

Including diverse groups of children and young people in health promotion and public health research

A review of methodology and practice

Report written by Theo Lorenc, Angela Harden, Ginny Brunton, Ann Oakley

EPPI-Centre
Social Science Research Unit
Institute of Education
University of London

EPPI-Centre report no. 1609 • September 2008

AUTHORS

Theo Lorenc, Angela Harden, Ginny Brunton, Ann Oakley of the EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

CONFLICTS OF INTEREST

There were no conflicts of interest for this review.

ACKNOWLEDGEMENTS

The authors would like to thank the members of the EPPI-Centre health team on whose work and experience they drew in preparing this report.

FUNDING

This work was undertaken by the EPPI-Centre, which received funding from the Department of Health. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health.

CONTRIBUTIONS

The protocol was developed by AH, TL, GB, and AO. The searching, mapping and data analysis procedures were undertaken by AH, TL and GB. The report was written by TL, AH and AO.

This report should be cited as: Lorenc T, Harden A, Brunton G, Oakley A (2008) *Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

The opinions expressed in this publication are not necessarily those of the EPPI-Centre or the funders. Responsibility for the views expressed remains solely with the authors.

ISBN: 978-0-9559087-4-3

© Copyright

Authors of the systematic reviews on the EPPI-Centre website (<http://eppi.ioe.ac.uk/>) hold the copyright for the text of their reviews. The EPPI-Centre owns the copyright for all material on the website it has developed, including the contents of the databases, manuals, and keywording and data-extraction systems. The centre and authors give permission for users of the site to display and print the contents of the site for their own non-commercial use, providing that the materials are not modified, copyright and other proprietary notices contained in the materials are retained, and the source of the material is cited clearly following the citation details provided. Otherwise users are not permitted to duplicate, reproduce, re-publish, distribute, or store material from this website without express written permission.

TABLE OF CONTENTS

SUMMARY	1
Background.....	1
Research questions	2
Findings	2
Conclusions	3
The contribution of this review.....	3
1. BACKGROUND	4
1.1 Diversity: rationales for inclusion	4
1.2 The UK policy context	4
1.3 Aims and objectives	5
2. METHODS USED IN THE REVIEW.....	7
2.1 Review design.....	7
2.2 Methodological and conceptual framework.....	7
2.3 Analysis of HP and PH research	9
2.4 Topic focus of the studies.....	10
2.5 Country of study	11
2.6 Year of publication	11
2.7 Study type and setting.....	11
3. ANALYSIS OF HP and PH RESEARCH: RESULTS	13
3.1 Target populations	13
3.2 Who exactly was included in the studies?	14
3.3 Conclusion	23
4. METHODOLOGICAL AND CONCEPTUAL FRAMEWORK.....	24
4.1 Introduction	24
4.2 Diversity and research: general issues.....	24
4.3 Stages of the research process.....	29
4.4 Conclusion	42
5. DISCUSSION AND RECOMMENDATIONS.....	43
5.1 Principal findings.....	43
5.2. Frameworks for good practice in research.....	45
5.3 Strengths and limitations of this review	46
5.4 Recommendations	47
6. REFERENCES.....	53

SUMMARY

This report presents the findings of a methodological study on the inclusion of diverse groups of children and young people in health promotion (HP) and public health (PH) research.

The need for diversity in research is increasingly recognised by researchers, funders and other stakeholders. Recent and ongoing changes to equalities legislation and health policy in the UK foreground the importance of the inclusion in research of diverse groups. An inclusive approach to research also improves the generalisability of research findings. However, there is relatively little guidance for researchers as to how diversity in research might be promoted. While valuable and relevant work has been done, few researchers have attempted to draw together material relevant to diversity across different populations and research designs. This report describes such an attempt.

We reviewed the methodological and conceptual literature on diversity, and identified ways in which it is appropriate for scientific and/or ethical reasons that HP and PH research should include socially diverse populations. They looked at 174 studies included in nine previous systematic reviews concerned with HP and PH research relating to children and young people. We found a mixed and patchy picture. Many studies do not provide sufficient socio-demographic detail to judge the nature of research samples. Rationales for sample selection are often unclear, and attention to diversity issues in data analysis appears to be driven more often by convenience than by science. Guidelines are suggested for improving the value and generalisability of HP and PH research – for all age groups as well as children and young people – by helping researchers to think critically about how social diversity can most appropriately be reflected in policy-relevant research.

Background

There are both scientific and ethical reasons why research should reflect the diverse and multicultural nature of many societies today. The inclusion of diverse groups in research increases the generalisability to the population at large of research findings, and helps to provide valuable information about differential patterns of ill-health and wellbeing. Apart from the benefit which people may gain directly from being involved in research, everyone has an equal right to contribute to research findings: the exclusion of certain groups is both unjust in itself and likely to contribute to injustice on the broader socio-political level. These rationales for the inclusion of diverse groups in research have been given added impetus in recent years in the UK by the development of a legal and policy agenda for diversity.

Previous studies have investigated the inclusion of diverse groups in research and found that this varies enormously. Poor reporting often makes it difficult to tell who was included. Most of this research has focused on clinical trials. Research using other methodologies, such as surveys and the many forms of qualitative research also needs to reflect diversity if it is to be reliable, ethically sound and relevant to policy and practice.

Research questions

This review seeks to answer the following two questions:

- What methods are appropriate for sampling, recruiting and retaining diverse groups of children and young people in HP and PH research, as well as for collecting and analysing data?
- To what extent have these methods been used in HP and PH research with children and young people?

The review was conducted in two parts. Firstly, theoretical and methodological literature was drawn upon to develop a framework for understanding diversity in research. Secondly, 174 studies included in nine previous reviews conducted by the EPPI-Centre were reanalysed to determine how HP and PH research has addressed diversity issues. The studies included in these nine reviews focused on HP and PH research related to mental health, healthy eating, physical activity, walking and cycling, accidental injury, and teenage pregnancy and parenthood. All the reviews were concerned with research involving children and young people. The studies included a range of research designs: 53 were evaluations of HP and PH interventions, using a controlled or randomised controlled trial design; and 121 studies examined people's views and experiences using a range of quantitative and qualitative methods (surveys, interviews, ethnographic and action research).

Findings

It was found that reporting of demographic data on research participants is frequently insufficient to determine whether or not studies included diverse groups of children and young people.

In around one in twenty studies, information was lacking on the number of participants or their age or gender; more than half provided no data on ethnicity, and more than two out of five gave no information about participants' socioeconomic position. Measures used (for instance relating to socioeconomic position) were often different between studies, making comparison of samples and findings difficult. Sampling strategies were unclear in most of the studies. A common claim, especially in the views studies, was that participants represented diverse groups, but data supporting this claim was often missing. More than one in ten of the studies did not specify the settings used to recruit participants, and over half did not mention whether or not consent was obtained. It was uncommon for the tools and measures used to have been validated with diverse groups. Subgroup analysis was carried out in half of the intervention studies, but, in most of these, no rationale was given. In almost half the qualitative studies which quoted material directly from interviews or focus groups with participants, no demographic descriptors were offered, and a similar proportion of studies gave no clear description of data analysis methods. In general, the intervention studies were better reported and paid more attention to diversity issues than the views studies.

Conclusions

Different areas of research vary widely in the extent to which they have investigated diversity-related questions, and it is difficult to establish a single set of methods or criteria that can capture the diversity-related questions relevant to research in the HP and PH field. The most appropriate means of achieving diversity will vary depending on the research questions or hypotheses, the methods employed, and the population and setting under study. At present, poor reporting makes it difficult to judge the inclusiveness or appropriateness to diversity issues of much HP and PH research.

A distinction needs to be drawn between research whose central questions focus on diversity, or on particular disadvantaged groups, and research whose questions relate to the general population. Much discussion of diversity in research to date has focused on the former, without considering implications for the latter. Including diverse groups in research means *both* conducting highly focused research with particular groups, to ensure that diverse voices are heard, *and* conducting more broadly framed research projects which genuinely reflect the diversity of the population.

This review identified areas in which research practice could improve. We recommend that it should be normal practice for all research to report major demographic factors for samples at baseline. In terms of strategies, to widen the research base, it may be useful for researchers to employ specific strategies such as, for example, testing sampling and recruitment strategies for their capacity to promote the inclusion of diverse groups; recruiting outside institutional settings, such as schools; and using data collection tools piloted with, or previously validated for, diverse populations.

More detailed reporting of sampling procedures (both of institutional settings, where relevant, and of individuals) would be valuable for all study designs, as would reporting of data analysis methods in qualitative research. More empirical work is needed on the effect of different research practices, particularly at the sampling and recruitment stages, on the diversity of samples obtained and the contribution of diverse groups to research findings. Other areas in which methods development would be valuable include sampling strategies for trials and qualitative research; how research synthesis, both quantitative and qualitative, can take account of diversity; the balance between foregrounding diversity and observing common themes in the collection and analysis of qualitative data; and the formation of research agendas and the research-policy relationship.

The contribution of this review

The work described in this report is a contribution to ongoing debates about the extent to which the social diversity of populations is, or might be, reflected in research, in order to improve the scientific credibility and policy relevance of research findings. The review focuses on HP and PH research related to children and young people, but many of the issues discussed here are relevant to social research more generally.

1. BACKGROUND

1.1 Diversity: rationales for inclusion

Diversity is usually considered across a number of intersecting dimensions, such as age, gender, ethnicity, religion, income, place of residence, occupation and sexual orientation (Department of Health 2005; Evans and Brown 2003). The implementation and evaluation of policy needs to take into account the characteristics and circumstances of the diverse social groups that are likely to make up any human population. There are also important scientific and ethical reasons to ensure that research reflects the diverse and multicultural nature of many societies today. An inclusive approach to research means that research findings are more generalisable to the population at large than would otherwise be the case. It also helps to provide valuable information about differential patterns of ill-health and wellbeing between groups (Bartlett et al. 2005, Flaskerud and Nyamathi 2000, Oakley 2006). Ethically, apart from the benefit which people may gain directly from being involved in research (Edwards et al. 1998), everyone has an equal right to participate in research and to contribute to research findings. Where the generalisability of research is compromised by its failure to include diverse groups, the implementation of policies or practices based on that research may lead to the further marginalisation of excluded groups, and thus contribute to injustice on the wider socio-political level (Baird 1999, Kahn et al. 1998).

Seeing diversity in terms of barriers and problems risks underestimating the positive value of an inclusive approach to research, both in terms of gaining more adequate answers to existing questions, and in shaping new fields of enquiry. Research has the potential actively to contribute *both* to the empowerment of disadvantaged or marginalised groups and to positive social change (Greenwood and Levin 1998, Kemmis and McTaggart 2005).

However, not a great deal is known about the extent to which diverse groups are included in research. This is true of research in general, and also of health-related research. Previous work has identified a substantial problem of non-reporting: many research reports simply give little information about the socio-demographic profile of research participants (Bartlett et al. 2005, Boehmer 2002, Ness et al. 1997, Oakley 2006, Oakley et al. 1998), so that it is impossible to judge to what extent research has taken account of diversity.

1.2 The UK policy context

The DH Research and Governance Framework for Health and Social Care states that research should, '[w]henever relevant ... take into account in its design, undertaking and reporting age, disability, gender, sexual orientation, race, culture and religion' so that 'the body of research evidence available to policy makers ... reflect[s] the diversity of the population' (Department of Health 2005, p 8). The equality duties for public bodies in the UK are currently outlined in three pieces of legislation on race relations, disability discrimination and gender equality. The Department of Health (DH) has made a strategic commitment to adopt a Single Equality Scheme (SES) which encompasses six 'strands of equality': race,

disability, gender, age, sexual orientation and religion and belief (Department of Health 2007).

These and other policy developments foreground the value of diversity and the role of public bodies in promoting it. One implication is a shift from a 'complaints-led' towards a 'proactive' and 'outcome-focused' view of anti-discrimination legislation (Department for Communities and Local Government 2007, Fredman and Spencer 2006). This revised view implies a move beyond guaranteeing equality of treatment, or equality of opportunity, to working actively in order to achieve equality of outcomes for diverse groups. The elimination of discrimination needs to be driven by a commitment in the form of action plans or equality schemes which are built into organisational structures, including the structures which support and drive research.

The need to understand diversity and inequality means an increased demand for a relevant evidence base. The broad shift in thinking represented by recent and ongoing policy changes forms a powerful framework for thinking about the relevance of diversity to research. However, such a framework does not provide detailed guidance on *when* or *how* to take into account 'age, disability, gender, sexual orientation, race, culture and religion', nor on whether all these aspects of diversity need to be addressed, or whether others are missing.

The review described in this report takes a close look at research in a particular area – health promotion and public health for children and young people. It examines how inclusive such research has been, and discusses ways in which it might better attend to diversity in its design, conduct and reporting. These findings should be useful in promoting discussion of how researchers and research commissioners can respond to the diversity agenda.

Health promotion (HP) and public health (PH) span an extremely broad field, covering a range of interventions and other sorts of activities that go beyond the boundaries of health care and treatment to promote health and prevent disease. Definitions of what exactly constitutes HP and PH have varied over time, especially in the relative emphasis given to individual and social determinants of health (Peersman 2001). The concept of 'social determinants' is an important one in contemporary definitions of HP and PH. This concept is concerned with individual risk factors, such as smoking and healthy eating as well as the 'social conditions that keep people well', such as housing, education and financial resources (Graham 2006, p 67). For the purposes of the present study, we included research on a wide range of topics and types of interventions including those designed to tackle the wider social determinants of health. All the research examined focused on whole populations, rather than on individual patients, and sought to understand the factors promoting health.

The focus throughout the report is on children and young people between the ages of 4 and 21 years, although many of the conclusions are applicable to all age groups. The issues arising from research with very young (preschool-aged) children are not specifically discussed; recommendations for good practice in research with this age group can be found, for example, in Alderson (2008).

1.3 Aims and objectives

The objectives of the work described in this report were as follows:

- to assess the extent to which HP and PH research has included diverse groups of children and young people with respect to dimensions such as age, gender, social class, sexual orientation, ethnicity, culture and religion
- to examine the ways in which HP and PH research with children and young people (aged 4-21 years) has attended to diversity in its design, sampling, data collection, and data analysis
- to assess to what extent researchers' rationales for considering or not considering diversity are appropriate to their research aims, setting and population, and to the context of the research
- to identify appropriate methods for sampling, recruiting and retaining diverse groups of children and young people in HP and PH research and for collecting and analysing data from these diverse groups
- to make recommendations for how research practice in this area might be improved with respect to including diverse groups of children and young people

We followed a two-stage process. Firstly, we developed a framework to identify relevant methodological and theoretical literature, and to clarify how HP and PH research could attend to diversity. Secondly, we analysed a set of primary studies included in reviews on various areas of HP and PH research, to investigate who is included in research and how the methods employed by researchers attend, or fail to attend, to diversity.

In this report, the results of the two stages are integrated to provide a broad overview of diversity in HP and PH research, and to contribute to the key debates. Chapter 2 describes the research methods used. Chapter 3 presents the empirical findings from analysis of HP and PH studies. Chapter 4 draws together findings from the methodological and conceptual framework, and from the analysis of primary studies. Chapter 5 discusses the main findings and the strengths and limitations of the review within the broader context of debates on diversity, and outlines points of reflection for researchers, including recommendations as to how HP and PH research might improve its attention to diversity issues.

The review presented in this report was undertaken at the same time as a parallel review, also drawing on previous EPPI-Centre HP and PH work, examining the extent to which HP and PH research has explicitly addressed health inequalities among children and young people (Oliver et al. 2008). Diversity and health inequalities are overlapping agendas, so researchers and policymakers interested in these areas will find it valuable to read both reports.

2. METHODS USED IN THE REVIEW

2.1 Review design

The review was conducted in two parts:

- 1) the development of a methodological and conceptual framework for considering diversity in research
- 2) an in-depth analysis of studies included in EPPI-Centre HP and PH systematic reviews focused on children and young people

The first part investigated what is considered to be 'best practice' in the inclusion of diverse populations in research. The second part covered two bodies of literature: trials of intervention effects and studies examining children's and young people's perspectives and experiences ('views' studies). These latter types of studies have used a range of methods, including large-scale surveys of attitudes and behaviour, and studies employing 'qualitative' methods such as in-depth interviews or focus groups. The results of the first part of the review informed the data-extraction tools developed for the second part.

2.2 Methodological and conceptual framework

The starting point in developing a methodological and conceptual framework to investigate diversity was twofold:

- 1) the six dimensions of diversity included in the Single Equality Scheme (Department of Health 2007): gender, age, ethnicity, religion, disability and sexual orientation (socioeconomic position (SEP) was added as a seventh dimension)
- 2) the stages of the research process: sampling, recruitment, data collection, and data analysis

These dimensions and stages provided a broad structure within which to examine the literature on diversity. We conducted preliminary searches on MEDLINE and the Cochrane Methodology Database, combining terms for 'children and young people' with terms for diversity (e.g. age, gender, ethnicity) and stages of the research process (e.g. sampling, data collection). These preliminary searches showed a lack of a discrete body of literature focused on including diverse groups of children and young people in HP and PH research, and the absence of any widely agreed conceptual framework for addressing questions of diversity as they relate to HP and PH research methods. In order to locate literature for the methodological and conceptual framework, we had to proceed in a pragmatic way, by identifying a number of bodies of literature to inform different areas of the framework:

- including women in clinical trials (e.g. Baird 1999)
- including ethnic minority groups in clinical trials (e.g. Swanson and Ward 1995)
- ethnicity, health and research (e.g. Nazroo 2006)
- disability, health and research (e.g. Gilson and DePoy 2004)
- sexuality, health and research (e.g. Boehmer 2002)
- research with children (e.g. Fraser et al 2004)
- research ethics, particularly for research with children (e.g. Morrow and Richards 1996)
- research on 'hard to reach groups' (e.g. those excluded from schools, refugees, asylum seekers) (e.g. Elliot et al. 2002, Jones and Allebone 1999)
- methods for sampling, data collection and data analysis in survey research (e.g. De Vaus 2002, Fowler 2002)
- methods for sampling, data collection and data analysis in qualitative research (e.g. Denzin and Lincoln 2005)
- quality assessment tools for qualitative research (e.g. Boulton and Fitzpatrick 1997; Spencer et al. 2003)
- justice and research (e.g. Kahn et al. 1998)
- existing policy discussion documents on diversity and research (e.g. National Co-ordinating Centre for Research Methodology 2001)

Papers were identified through an iterative process including personal contact, Google Scholar, and 'snowballing' methods, such as reference-scanning and searches on citation indexes (Science Citation Index and Social Science Citation Index). Our goal in identifying reports was not to cover all these bodies of literature comprehensively, but rather to identify key papers in each field and summarise the key theoretical and methodological issues, in order to generate a set of hypotheses as to how researchers could attend to diversity in HP and PH research with children and young people.

These search strategies reflect those adopted in the literature on reviewing complex evidence (Dixon-Woods et al. 2006, Greenhalgh and Peacock 2005). The idea of 'theoretical saturation' found in some of these literatures, for example, guided the process of searching used: although saturation in the strong sense is probably too ambitious a goal where the conceptual ground covered is very extensive, decisions as to how much searching was sufficient were guided by a sense of relative closure of certain areas of the conceptual framework relative to others. The attempt to integrate this material into an overarching framework cannot be seen as a 'synthesis' in the narrow sense or even as a 'synthetic construct' in rather broader sense of Dixon-Woods et al. (2006), since it does not stand over and above the included reports as part of a distinct order of explanation, but rather seeks to advance and shape an emerging dialogue between them.

2.3 Analysis of HP and PH research

For the empirical part of the review, we examined researchers' treatment of diversity in more depth by interrogating a dataset of 174 specific HP and PH studies. These were drawn from nine previous EPPI-Centre HP and PH reviews in the topic areas of accidental injury, mental health, healthy eating, physical activity, walking and cycling and teenage pregnancy and parenthood (Table 2.1). Studies from these reviews were selected because they had already been retrieved and coded, and so formed a conveniently available body of research. The reviews covered a broad range of topics of health promotion and public health work with children and young people. They were all based on extensive searching and systematic inclusion procedures.

Table 2.1: Source of studies in the dataset (N = 174)

Review title	Intervention studies N (%)	Views studies N (%)
Young people and mental health: a systematic review of barriers and facilitators (Harden et al. 2001)	5 (9%)	12 (10%)
Young people and physical activity: a systematic review of barriers and facilitators (Rees et al. 2001)	0	12 (10%)
Young people and healthy eating: a systematic review of barriers and facilitators (Shepherd et al. 2001)	3 (6%)	7 (6%)
Children and physical activity: a systematic review of barriers and facilitators (Brunton et al. 2003)	3 (6%)	3 (2%)
Children and healthy eating: a systematic review of barriers and facilitators (Thomas et al. 2003)	18 (34%)	8 (7%)
A synthesis of research addressing children's, young people's and parents' views of walking and cycling for transport (Brunton et al. 2006)	0	30 (25%)
Young people, pregnancy and social exclusion: a systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support (Harden et al. 2006)	16 (30%)	22 (18%)
Accidental injury, risk-taking behaviour and the social circumstances in which young people live: a systematic review (Thomas et al. 2007)	0	24 (20%)
Studies included in more than one review	8 (15%)	3 (2%)
Total	53 (100%)	121 (100%)

The 174 studies included both evaluations of interventions and studies of children and young people's views and experiences. Seven of the original reviews included both types of studies, while two (accidental injury and walking and cycling) only included views studies. Included in this dataset of 174 studies were all the views studies in the original reviews, except for four which did not elicit views from the under-22 age group, and one which could not be sourced for this review. Views studies were included in the dataset of 174 studies, irrespective of

their methodological quality, since there are no widely agreed guidelines or quality assessment criteria for such studies. Intervention studies were only included if they had been judged methodologically sound in the original reviews, since there is more consensus in this area (Moher et al. 2001).

In the original reviews, detailed data had been collected from these studies about sample characteristics, study design, sampling and recruitment methods, and methods of data collection and analysis. We reanalysed the data and collected supplementary data using the full texts of the studies. The new analysis aimed to determine first, who was included in the samples of the studies in the dataset, and, second, to what extent the studies had used methods identified in our conceptual review as promising means to promote diversity.

2.4 Topic focus of the studies

Tables 2.2 and 2.3 show the topic focus of the 174 studies.

The topics obviously reflect the areas in which the original reviews were carried out, although some topics (e.g. alcohol, tobacco) were included in a number of studies without being a focus of any of the reviews. Healthy eating dominated the intervention studies (N=26, 49%), and physical activity the views studies (N=49, 40%).

Table 2.2: Topic focus of the intervention studies (N = 53; some studies focused on more than one topic)

	N (%)
Healthy eating	26 (49%)
Pregnancy prevention / teenage parenthood	14 (26%)
Physical activity	11 (21%)
Mental health	6 (11%)
Tobacco	4 (8%)
Obesity	3 (6%)

Table 2.3: Topic focus of the views studies (N = 121; some studies focused on more than one topic)

	N (%)
Physical activity	49 (40%)
Accidents / injury	22 (18%)
Pregnancy prevention / teenage parenthood	22 (18%)
Healthy eating	18 (15%)
Mental health	15 (12%)
Alcohol	7 (6%)
Drugs	5 (4%)

2.5 Country of study

Most of the intervention studies were conducted in the USA (N=43, 81%). Eight (15%) were conducted in the UK, and one each in Finland and Norway.

All the views studies were UK studies. One compared the views of children in the UK and Germany (Neale et al. 1998).

2.6 Year of publication

Most of the studies were published after the mid-1990s; 64% of the intervention studies and 85% of the views studies were dated 1995 or later.

2.7 Study type and setting

Of the intervention studies, 16 were non-randomised controlled trials and 37 were randomised controlled trials (RCTs). The most common settings for the interventions were primary or secondary education, the home and the community. The largest group consisted of educational interventions delivered by teachers in school settings (N=32, 60%). Many of the interventions in this group also included other components (e.g. changes in school meal provision or parent-delivered home-based interventions). Ten studies evaluated multi-component interventions aimed at improving outcomes for young parents or disadvantaged young people at risk of pregnancy. The interventions evaluated in the remaining studies were either based in health services or in community sites, or were non-educational interventions (e.g. service provision) delivered in schools.

The views studies adopted a range of methodologies, from large-scale surveys, using closed questionnaires and predominantly quantitative methods, to ethnographic and action-research studies with an emphasis on in-depth interpretive understanding. Nearly half the studies were in-depth interviews and/or focus group studies, and nearly one-third were large-scale surveys of attitudes and behaviour using self-completion questionnaires; the remainder combined the

two approaches. The majority of the views studies (N=69, 57%) were also conducted wholly or mostly in primary or secondary educational settings.

The next chapter presents an assessment of the extent to which diversity was incorporated as an element in the design, analysis and reporting of the 174 studies.

3. ANALYSIS OF HEALTH PROMOTION AND PUBLIC HEALTH RESEARCH: RESULTS

In this chapter, we present the findings from our reanalysis of the 174 studies included in previous EPPI-Centre health promotion (HP) and public health (PH) reviews.

Most of the analyses are presented separately for the intervention studies (N=53) and the views studies (N=121).

3.1 Target populations

Of the 53 intervention studies, 17 (32%) explicitly stated that their aim was to target disadvantaged groups. All but two of these concerned teenage pregnancy and parenthood, and targeted either teenage parents or groups considered to be at risk of teenage pregnancy. The other 36 studies (68%) were carried out with a general population sample. Of the views studies, 32 (27%) targeted specific populations, and 91 (75%) studied the population of children and young people as a whole; these figures include two studies which both sampled from the general population and focused on specific groups within this population.

Table 3.1 shows the target populations in the interventions and views studies; some populations fall in more than one category – for example, teenage mothers are counted both as ‘women’ and as ‘young parents’.

Table 3.1: Populations targeted in the intervention (N = 53) and views studies (N = 121; some studies targeted more than one population)

	Intervention studies N (%)	Views studies N (%)
General population	36 (68%)	91 (75%)
Young parents / ‘At risk’ of teenage pregnancy	15 (28%)	22 (18%)
Girls / Young women	3 (6%)	19 (16%)
Boys / Young men	0	3 (2%)
Looked-after / In local authority care	0	3 (2%)
Ethnic minority	3 (6%)	0
Homeless	0	2 (2%)
With mental health problems	0	1 (1%)
Low-income	0	1 (1%)
Young offenders	0	1 (1%)

Of the views studies, in addition to the 32 studies which targeted specific groups, a further 19 studies stated that one of their primary aims was to investigate diversity by comparing different subgroups within the sample or by exploring the influence of demographic variables. In 10 of these studies the axis of diversity examined was gender; in six, it was age; in four, socioeconomic position (SEP); and, in three, ethnicity. No intervention studies reported an explicit primary aim of comparing subgroups within the sample.

3.2 Who exactly was included in the studies?

3.2.1 Reporting of demographic information

Studies varied widely in how much information they reported about the demographic characteristics of their sample. Some studies did not provide information which was sufficiently precise to be aggregated in this way. For the analyses below, studies were required to (a) report information about participants themselves, and not, for example, about the community or school from which they were recruited; (b) report distributions or ranges of the relevant variable for the whole sample, not for a part of the sample only; and (c) report demographic information in the form of a precise characterisation of the individuals in the sample, not only a broad qualitative description of the sample as a whole. Tables 3.2 and 3.3 show the intervention and the views studies separately. The first column in each table shows the number and percentage of studies which provided no usable information on sample characteristics; the second column shows those studies from which some information could be inferred, no matter how vague or at what level; and the third shows those studies which gave sufficiently precise information to be included in the analysis.

Table 3.2: Reporting of demographic information in the intervention studies (N = 53)

	No information N (%)	Some information, but not complete N (%)	Complete information for all participants N (%)
Number of participants	0 (0%)	2 (4%)	51 (96%)
Age	5 (9%)	10 (19%)	38 (72%)
Gender	1 (2%)	14 (26%)	38 (72%)
SEP	18 (34%)	22 (42%)	13 (24%)
Ethnicity	13 (24%)	4 (8%)	36 (68%)
Religion	53 (100%)	0 (0%)	0 (0%)
Sexual orientation	53 (100%)	0 (0%)	0 (0%)
Disability	53 (100%)	0 (0%)	0 (0%)

Table 3.3: Reporting of demographic information in the views studies (N = 121)

	No information N (%)	Some information, but not complete N (%)	Complete information for all participants N (%)
Number of participants	8 (7%)	6 (5%)	107 (88%)
Age	5 (4%)	26 (22%)	90 (74%)
Gender	11 (9%)	33 (27%)	77 (64%)
SEP	58 (48%)	39 (32%)	24 (20%)
Ethnicity	83 (69%)	16 (13%)	22 (18%)
Religion	117 (97%)	1 (1%)	3 (2%)
Sexual orientation	120 (99%)	0 (0%)	1 (1%)
Disability	119 (98%)	2 (2%)	0 (0%)

3.2.2 Number of participants

Of the 174 studies, eight (5%) provided no information on number of participants, and a further eight (5%) gave some, but not exact, information.

Intervention studies

Only two (4%) of the 53 intervention studies did not report the exact number of participants. The remaining 51 studies represent a total of 59,901 participants. Sample sizes ranged from 28 to 9,685, with a mean sample size of 1,175 and a median of 628. (The total N for each study has been calculated as the number of participants for whom at least some baseline data was available, and includes both intervention and control groups.)

Views studies

Fourteen (12%) of the 121 views studies did not clearly report the total number of participants. In eight (7%), no information was reported, while six (5%) reported some information but not the exact total. The 107 studies which did report the number of participants represent a total of 85,853 participants. Sample sizes ranged from 3 to 18,221, with a mean sample size of 802 and a median of 160. (Where studies included two or more stages of sampling – such as a large survey followed by in-depth interviews with a subsample of survey respondents – the largest Ns have been chosen in each case.)

3.2.3 Age

Intervention studies

Five studies (9%) did not report the age of participants, and ten (19%) reported incomplete information, such as an overall mean only. The remaining 38 studies, representing a total of 28,171 participants, reported either age ranges for the total sample or exact distributions. The mean age was 12.1 years. The aggregate age distribution for the samples in the intervention studies (assuming the sample to have been evenly distributed across the age range except where exact distributions were given) is shown in Table 3.4.

Table 3.4: Ages of participants in the intervention studies (N = 28,171)

Age	N (%)
<4	114 (0.4%)
4	88 (0.3%)
5	176 (0.6%)
6	279 (1.0%)
7	821 (2.9%)
8	2,683 (9.5%)
9	4,009 (14.2%)
10	2,442 (8.7%)
11	2,183 (7.7%)
12	2,532 (9.0%)
13	3,353 (11.9%)
14	2,448 (8.7%)
15	2,949 (10.5%)
16	1,079 (3.8%)
17	828 (2.9%)
18	722 (2.6%)
19	587 (2.0%)
20	439 (1.6%)
21	439 (1.6%)
>21	0 (0%)

Views studies

Five studies (4%) did not report the age of participants, and 26 (21%) reported incomplete information, such as an overall mean only. The remaining 90 studies, representing a total of 75,565 participants, reported either age ranges for the total sample or exact distributions. The mean age for the views studies was 13.6 years. The aggregate age distribution for the samples in the views studies (assuming the sample to have been evenly distributed across the age range except where exact distributions were given) is shown in Table 3.5.

Table 3.5: Ages of participants in the views studies (N = 75,565)

Age	N (%)
<4	93 (0.1%)
4	163 (0.2%)
5	243 (0.3%)
6	258 (0.3%)
7	300 (0.4%)
8	322 (0.4%)
9	1,308 (1.7%)
10	1,540 (2.0%)
11	5,557 (7.4%)
12	9,670 (12.8%)
13	12,438 (16.5%)
14	17,941 (23.7%)
15	15,909 (21.1%)
16	5,273 (7.0%)
17	1,525 (2.0%)
18	1,148 (1.5%)
19	445 (0.6%)
20	426 (0.6%)
21	345 (0.5%)
>21	661 (0.9%)

3.2.4 Gender

Intervention studies

One study (2%) reported no information on the gender of study participants. Fourteen (26%) reported some information on gender: for example, that the sample contained both males and females, but failed to give the number of males and females in the sample. The 38 studies for which gender was reported covered a total of 50,773 participants, comprising 15,308 males (30%) and 35,465 females (70%). Ten of these studies sampled only or predominantly females (of which nine focused on teenage pregnancy). If these studies are removed, the remaining 27 studies included 14,961 males (49%) and 15,281 females (51%).

Views studies

Eleven studies (9%) reported no information on the gender of study participants. Thirty-three (27%) reported some information on gender: for example, that the sample contained both males and females, but failed to give the number of males and females in the sample. The remaining 77 studies included a total of 31,148 females (57%) and 23,445 males (43%). Of these studies, 25 sampled wholly or predominantly single-gender groups. If these studies are removed, the remaining 52 are equally balanced, with 22,951 females (50%) and 23,053 males (50%).

3.2.5 SEP

There was no information on the SEP of the sample in 76 (44%) of the studies. Sixty-one studies (35%) gave some information: studies in this category either gave qualitative information on the SEP of the sample, but no quantitative measure, and/or only information on the average SEP of participants' schools or communities, and not on those of participants themselves. In 37 studies (21%), there was more precise quantitative information reported either for participants themselves, or for their parents or their families as a whole. Measures of SEP in these 37 studies were quite heterogeneous, as shown in Table 3.6. Welfare receipt and occupation were the most commonly used measures (each was used in 12 studies).

Table 3.6: Measures of SEP used (N = 37 studies reporting this information; some used more than one)

	N (%)
Welfare receipt	12 (32%)
Occupation	12 (32%)
Education	8 (22%)
Unemployment or employment type (e.g. fulltime, part-time)	6 (16%)
Household income	4 (11%)
Car ownership	2 (5%)
Housing type	2 (5%)

Intervention studies

Thirteen of the intervention studies (25%) characterised the SEP of the individuals in the sample precisely and are included in Table 3.6. A further seven studies used phrases such as ‘middle class’ or ‘disadvantaged’, but used no objective measure to characterise the SEP of the sample. Six studies described the average SEP of the schools from which the sample was drawn, using the percentage of pupils eligible for free school meals, and one used median household income to characterise the SEP of the community among which the intervention was carried out. A further eight studies gave only a qualitative indication of the SEP of either the school or the community.

Views studies

Twenty-four views studies (20%) provided some quantitative measure of the SEP of the individuals in the sample, and are included in Table 3.6. Eleven of these used an occupational class measure of SEP. Five of these 11 studies divided samples into two classes (‘middle class’ vs. ‘working class’ or ‘non-manual’ vs. ‘manual’), three studies used three classes (AB / C1C2 / DE), and three used four (AB / C1 / C2 / DE). Thirteen studies (11%) gave quantitative information on the SEP either of the schools or of the communities from which samples were drawn. For schools, the measure used was the percentage of pupils receiving free school meals (N=5). For communities, measures included unemployment rate (N=5), car ownership (N=2) and average household income (N=1). Six studies used compound measures of community-level deprivation, such as the Carstairs index or the Index of Multiple Deprivation. Twenty-five studies (21%) gave some qualitative information about the social class of schools or communities, or stated their intention to sample across a range of social classes, but did not provide clear quantitative data.

3.2.6 Ethnicity

There was no information on the ethnicity of samples in 96 (55%) of the 174 studies; 20 (11%) gave some information and 58 (33%) provided detailed information.

Intervention studies

Thirty-six of the 53 intervention studies (68%) provided information on the ethnicity of participants. Data on ethnicity were available for a total of 47,782 participants (Table 3.7). The two largest categories were 'White' (43% of participants) and 'Black' or 'African-American' (40% of participants). All but two of these 36 studies were carried out in the USA.

Table 3.7: Ethnicity descriptors in the 36 intervention studies which provided information (N = 47,782 participants)

	N (%)
'White'	20,680 (43%)
Non-white, 'other' or 'ethnic minority', not further specified	1,038 (2%)
Black / African-American	19,253 (40%)
Hispanic	6,183 (13%)
Asian	381 (<1%)
Native American	38 (<1%)
Mixed	209 (<1%)
Total	47,782 (100%)

Views studies

Twenty-two views studies (18%) provided information on the ethnicity of participants. Data on ethnicity was available for a total of 8,945 participants (Table 3.8). Most (81%) participants were described as 'White'; 7% of the studies referred to 'non-White' or 'ethnic minority' samples without being more specific.

Table 3.8: Ethnicity descriptors in the 22 views studies which provided information (N = 8,945 participants)

	N (%)
'White'	7,223 (81%)
Non-white, 'other' or 'ethnic minority', not further specified	593 (7%)
Asian: Pakistani	20 (<1%)
Bangladeshi	47 (<1%)
Indian	5 (<1%)
Other / Not specified	711 (8%)
Black: Black African	60 (<1%)
Black Caribbean	79 (<1%)
Other / Not specified	109 (1%)
Chinese	3 (<1%)
Middle Eastern	1 (<1%)
Mixed	94 (1%)
Total	8,945 (100%)

3.2.7 Religion

Only four studies gave information on participants' religion, and only three reported precise distributions (Table 3.9).

Intervention studies

No intervention studies provided information on the religion of participants.

Views studies

Three views studies reported precise information on participants' religion; the data covered 369 participants.

Table 3.9: Descriptors of religion in the three views studies which provided information (N = 369 participants)

	N (%)
No religion	93 (25%)
Christian: Church of England	143 (39%)
Catholic	58 (16%)
Other / Not further specified	34 (9%)
Muslim	33 (9%)
Jewish	1 (<1%)
Sikh	1 (<1%)
Hindu	1 (<1%)
Other	5 (1%)
Total	369 (100%)

3.2.8 Disability

One study listed disabled participants in the category of 'employment status', and one gave the number of participants with 'special needs'; neither gave detailed information.

3.2.9 Sexual orientation

One study, of drug use among young people in nightclubs, reported the sexual orientation of participants.

3.3 Conclusion

The results of our analysis of 174 HP and PH studies relating to children and young people show that reporting of socio-demographic characteristics of the samples included in research is frequently insufficient to evaluate whether or not diverse populations have been included. Particularly striking is the lack of information related to SEP (where over two in five of the studies gave no information) and ethnicity (more than half). However, information relating to age, gender or number of participants is also lacking in around one in 20 of the studies. The measures used (for instance related to SEP) are often different between studies, making comparison of samples and findings difficult. In general, reporting in the intervention studies is better than in the views studies.

4. METHODOLOGICAL AND CONCEPTUAL FRAMEWORK

4.1 Introduction

This chapter discusses the review of debates about how diversity might be represented in health promotion (HP) and public health (PH) research. The discussion is divided into two sections. The first section relates to the research process as a whole, and the second covers more specific issues, such as sampling, recruitment, data collection and data analysis. The research question and design are considered throughout the chapter.

Not all the issues investigated here are relevant to all research projects. While the broad issues are independent of research type or method and apply to HP and PH (and, indeed, much other) research in general, how they apply to particular research projects will vary depending on the research question and its resultant design, the context of the research and the specific methods employed. One important distinction is between studies whose research aims demand a specific focus on diversity issues, and those which do not have such a focus. The former include, for example, studies whose research aim includes a comparison of diverse groups with one another; those which aim for an understanding of the experiences of particular disadvantaged groups; and/or studies of topic areas which impact disproportionately on diverse groups. Studies without such a focus include those in topic areas which are not known to be strongly associated with any dimension of diversity. Different levels of engagement with diversity issues are appropriate in these two groups of studies.

4.2 Diversity and research: general issues

4.2.1 Defining diversity

The legal and policy framework of the diversity agenda identifies six dimensions of diversity or 'strands of equality' which should be attended to age, gender, ethnicity, religion, sexual orientation and disability (Department of Health 2007). However, other dimensions of diversity are also relevant, including both broad demographic categories (socioeconomic position (SEP), employment status, social capital) and particular risks of disadvantage (such as those involved in the status of refugee, asylum seekers, non-English speakers, travellers, prisoners, children and young people excluded from school, and those with long-term health problems). Depending on the context of the research and the research question, some of these dimensions may not be relevant, while others may be important foci of investigation. Since, in many cases, different forms of disadvantage are strongly correlated both with each other and with negative health outcomes, many different dimensions of diversity may be relevant to a given research question. The challenge which this poses for researchers is that, for most questions in the HP and PH field, obtaining the most relevant data will require going beyond the populations which are easiest to access and the methods which are easiest to implement.

There are two broad approaches in the literature to the question of how to identify the relevant dimensions of diversity. In one, a standardised checklist of population variables (such as the six statutory dimensions) is used. In the other, diversity is defined differently in different contexts, depending on the aspects of diversity that are relevant for particular populations and topic areas. The use of standardised lists can have some advantages over a wholly flexible approach to defining diversity, but these may also detract attention from the need for researchers to reflect critically on the extent to which an inclusive approach to diversity flows from their particular research question. For example, does a study of an intervention for promoting healthy eating among schoolchildren require a SEP or gender-balanced sample? Should the results be analysed according to SEP and gender? Different issues relating to diversity are salient at different stages of the research process: the issue of sample representativeness as an aid to the generalisability of research findings, for example, is different from the question of the extent to which *a priori* hypotheses mandate subgroup analyses by variables such as SEP or ethnicity.

A further complication of defining diversity is that researchers' judgements as to which dimensions of diversity are relevant may conflict with research participants' perspectives. An example of this conflict can be found in one of the studies in the views dataset (Hall et al. 2003). This was a peer research project led by young people, who made most of the substantive decisions regarding data collection and analysis. One of these decisions was not to collect data on ethnicity or disability, 'because the peer researchers, who played a key role in designing the questionnaire, felt strongly that we should not collect this data, as they believed that the experience of teenage parenthood was the most important indicator within the survey' (p 7) – a decision which went against the wishes of the lead academic researcher on the project (p 34). In this study, then, a view of diversity as based on the empowerment of the population under study, and on the participation of this population in shaping the research question, came into conflict with a view of diversity as an *a priori* list of categories.

4.2.2 Measuring diversity

The question of how particular dimensions of diversity should be measured raises analogous concerns: should researchers use standardised categories to facilitate the monitoring of diversity and the comparison of samples and results across studies, or allow research participants flexibility in identifying the categories most relevant to them?

Many dimensions of diversity can be seen either in somatic-biomedical terms or as social identities; in some cases, the two can be distinguished terminologically (for example, somatic 'sex' as against social 'gender', and researchers should be careful not inadvertently to over-emphasise the former at the expense of the latter (Hussain-Gambles et al. 2004, Karlsen and Nazroo 2002, Lippman 2006). Ethnicity, for example, is a complex concept, including national origin, family background, cultural identity, and socio-political contexts. Individuals' identification with a particular ethnic identity, as well as their views on whether this identity is relevant to a particular question, may be context-sensitive (Gunaratnam 2003).

Disability as a dimension of diversity raises particular issues in health-related research. Standardised measures of disability, such as the World Health Organization's International Classification of Functioning, Disability and Health (<http://www.who.int/classifications/icf/en/>) and the Disability Index of the Health

Assessment Questionnaire (<http://aramis.stanford.edu/>), are geared towards measuring disability as an outcome rather than as a determinant or a demographic variable, and hence may be of limited value in promoting diversity and inclusion. Seeing disability as part of the diversity agenda implies attention to its social and cultural meanings, and its potential role as a pretext for discrimination, and these are not well captured by biomedical or impairment-based concepts (Gilson and DePoy 2004).

4.2.3 Language barriers

A key domain of concern with respect to the representation of cultural diversity in research is the barrier to inclusiveness posed by the use of the dominant language in research. In the UK and the USA, most researchers design their projects, recruitment procedures, measures and data collection tools in the English language. This has the immediate effect of disenfranchising those groups who are not fluent or are not literate in English, or for whom it is a second language. There is no reliable national data on language use in the UK (Aspinall 2007). A survey in the mid-1990s showed that between 16% and 44% of minority ethnic groups (Bengalis, Gujeratis, Punjabis, Chinese and refugees (Bosnians, Somalis, Tamils and Kurds)) did not have a level of 'survival competence' in written and spoken English (Carr-Hill et al. 1996).

In this dataset of 174 HP and PH studies – despite many studies being carried out with multi-ethnic populations – only two studies (Mauthner et al. 1993, Rogers et al. 1997) explicitly stated that translation or interpretation formed a part of the sample recruitment or data collection process. A further study (Armstrong et al. 1998) stated that information and consent forms were translated, but not data-collection tools. Translation may create its own problems. Key concepts used to identify outcomes of interest, for example, may not be adequately translatable into participants' native languages, or different translations may be possible (Grewal and Ritchie 2006). There is also a risk that the translation of qualitative data for presentation in research reports may distort the significance of the findings (Temple 2005). There are a number of reasons why research materials are not translated into minority languages and why recruitment procedures are not more linguistically inclusive: it may not occur to researchers that this is an issue; they may be concerned about ethics and cultural sensitivity; they may refer to the rationale of needing 'homogeneous' research samples; or budgets may be insufficient to fund translation/interpretation activities. This is an area of research methods which appears to have received little attention. An informative exception is a trial of a social support intervention for disadvantaged mothers carried out in London. Oakley and colleagues (2003) translated research materials into six languages and used interpreters speaking 25 different languages to help in the recruitment process. The result was balanced recruitment rates among English and non-English speakers. Employing interpreters resulted in a more socially disadvantaged sample: 85% of the women recruited using interpreters were socially disadvantaged compared to 51% of those recruited without using interpreters. Costs were 70% higher for interpreter-aided recruitment. Oakley and colleagues point out that most research in multicultural societies suffers from serious inattention to the impact on research samples and findings of practices which exclude minority groups.

4.2.4 Reporting diversity

The reporting of baseline demographic variables is still not widely regarded as part of good practice in research. The CONSORT guidelines for clinical trials, for example (Moher et al. 2001), do not recommend that such information be routinely reported. The International Committee of Medical Journal Editors recommends that, if variables such as age, sex or ethnicity form part of the inclusion criteria of the study, they should be reported on and their use explained, but not otherwise: 'Because the relevance of such variables as age and sex to the object of research is not always clear, authors should explain their use when they are included in a study report; for example, authors should explain why only subjects of certain ages were included or why women were excluded. The guiding principle should be clarity about how and why a study was done in a particular way. When authors use variables such as race or ethnicity, they should define how they measured the variables and justify their relevance' (www.icmje.org, section IV.A.6.a). However, the *British Medical Journal's* guidelines for authors require that authors report 'numbers entering and completing the study, sex, and ethnic group if appropriate' (<http://resources.bmj.com/bmj/authors/types-of-article/research>); regarding ethnicity, they further note: 'Ethnicity and culture are socially determined variables of limited use in biological research, though they are useful in health services research. All the variables are confounded by socioeconomic status' (<http://resources.bmj.com/bmj/authors/article-submission/article-requirements>).

Social science journals are generally less clear about how and why data relating to diversity should be reported. In a study by Newman and Elbourne (2004) of educational research, eight out of 11 educational research journals provided no such guidance as to how studies should be reported. Quality assessment tools for qualitative research often recommend that the sample should be described, but only sometimes do they specify that factors such as age, gender, ethnicity and SEP should be used to characterise the sample (Boulton and Fitzpatrick 1997, Boulton et al. 1996).

4.2.5 External validity and generalisation

One of the most important scientific rationales for including diverse groups in research is to ensure that research is externally valid – that its findings can be generalised to populations beyond the sample participating in the research. There are long-standing debates in the methods literature as to how this notion of external validity or generalisability should be understood, and, in the case of certain qualitative methodologies, as to whether it is appropriate at all.

The literature on external validity of trials of biomedical interventions has drawn attention to a number of factors which may reduce external validity. These include trial setting, characteristics of the intervention, and choice of outcome measures (Rothwell 2005a). The most significant issue from a diversity perspective is the mismatch between research samples and the populations which receive interventions in practice (Bartlett et al. 2005, Rothwell 2005a). In particular, women, ethnic minorities and older people are often not represented in clinical research in proportion to their numbers in the population as a whole (Bartlett et al. 2005, Hall 1999, Mason et al. 2003). This places severe limitations on how far the findings of such research can be generalised to the wider population, and can sometimes result in mistaken conclusions being drawn about the effects on subgroups of particular health care practices.

The analyses of baseline demographics in the dataset of HP and PH studies in Chapter 3 demonstrate that reporting is often insufficient to judge the external validity of studies. If it is simply not clear who has been included in a study, it is impossible to decide whether or not the findings of that study are relevant to the wider population. This is true of both the intervention and the views studies in this dataset. However, in the literature concerned with the methodology of qualitative research, there is an argument that the notion of 'generalisation' which is appropriate to qualitative research is distinct from, and broader than, that associated with quantitative studies, such as trials. According to this argument, the generalisability of qualitative findings is not impeded by the failure of research samples to be demographically representative of the population. Numerical representativeness is often not regarded as a relevant goal in sampling for qualitative research.

Some qualitative methodologists go beyond this to reject the notion of generalisability altogether, and replace it with a broader concept of 'transferability'. They argue that 'the burden of proof [that a particular result is transferable] lies less with the original investigator than with the person seeking to make an application elsewhere' (Lincoln and Guba 1985, p 298). The problem with this argument is that the interpretation of the findings of qualitative studies requires that they be placed in some context broader than the particular sample studied. If relevant information about who was included in the sample is not reported, the extent to which a given result is in fact transferable to other populations cannot be determined (Britten et al. 1995). In addition, many qualitative researchers implicitly or explicitly generalise their findings to populations much wider than the research sample.

4.2.6 Involvement of research participants

Involving research participants in making decisions about research is of value in promoting the inclusion of diverse groups throughout the research process. For example, involving ethnic minority communities in research is recognised as an important means of ensuring 'community legitimacy', promoting trust in researchers and, hence, increasing minority recruitment and encouraging full participation (Corbett et al. 1995, Nickens 1990). Involving the population also helps to ensure that research is relevant and sensitively conducted.

Johnson (2006) provides an overview of the benefits and costs of involving diverse groups, and of relevant strategies of involvement. He recommends 'social action research' as a model of good practice, incorporating the involvement of diverse groups in research design, data collection, and feedback on findings. Other recommendations developed in the involvement literature are also relevant to diversity issues. For example, recommendations for the involvement of children and young people in research include the following: 'transparency and agreement about which children (and why) will be the partners in the research; ... [ensuring] that research tasks that are identified are appropriate and 'do-able' and take into account children's views, age and abilities, as well as relevant social and cultural factors; that the language, methods and processes of research are made accessible to children; ... [and that] there are adequate support systems in place' (Jones 2004, p 129).

Another important benefit of involvement is to further the capacity of research participants to challenge existing assumptions. 'It may also be possible to include consumers or other stakeholders in panels with the primary responsibility of

questioning assumptions that are made about values, rather than representing the values of any particular group' (Schünemann et al. 2006). This idea has been seen by its proponents as something quite separate from the 'representation' of diverse groups. However, the emphasis on questioning accepted values is also useful in understanding the contribution that the involvement of diverse groups can make to research.

4.3 Stages of the research process

4.3.1 Sampling

The sampling and recruitment stages of the research process are of fundamental importance in the question of how research can promote diversity. For research to address diversity, it must sample and recruit participants from sufficiently diverse groups. The central concern is that sampling procedures should not exclude any subgroups of the population under study.

4.3.2 Exclusion criteria

One potential barrier to diversity is the adoption of exclusion criteria which are unduly restrictive. In health care randomised controlled trials (RCTs), criteria for exclusion of participants are often not clearly reported, and the rationale for excluding certain groups (e.g. women and older people) is weak or non-existent (Britton et al. 1999, Oakley 2006, Prescott et al. 1999, Van Spall et al. 2007). The exclusion of people with co-morbidities may also bias the sample in favour of younger and healthier participants. It is unclear whether such deliberate restrictions of the sample are a problem with HP and PH studies, which often seek to sample general populations without specific health needs. In this dataset of 174 studies, only three reported details of pre-specified exclusion criteria, and all gave some rationale for these criteria (Arai 2004, Epstein et al. 2001, Hughes et al. 1999). In addition, a number of intervention studies evaluated interventions which had their own eligibility criteria distinct from those of the study itself: for example, in studies of alterations to welfare systems, only people on a low income and resident in a particular area may have been eligible.

4.3.3 Sampling strategies

Two broad types of research question correspond to two types of sampling strategy. For research such as surveys, which aim to quantify attitudes or behaviour across a population, the primary concern is that the sample be numerically representative of the population, so that findings can be generalised from the former to the latter. For more in-depth and qualitative research which aims for rich description and analysis of individuals' beliefs, motivations or experiences, the primary aim is to obtain a sample which will provide the most illuminating data for a particular research issue. The sampling strategies corresponding to these two distinct research goals are sometimes referred to as 'statistical' and 'theoretical' sampling respectively (e.g. Flick 2006). Statistical sampling aims for a sample which represents the population under study, in the specific sense that all subgroups are present in the sample proportionally to their size in the population. Theoretical sampling aims for a sample which includes all

cases or perspectives (or as many as is reasonably practicable) which are relevant to the research question.

The most straightforward form of statistical sampling is simple random sampling, where cases are chosen at random from a pre-defined sampling frame, such as a list of names. As a further safeguard of the diversity of the sample, 'stratified' random sampling may be used, in which the sample is drawn randomly from sub-populations or 'strata' (e.g. different ethnic or SEP groups), to ensure that these subgroups are selected in the same proportion that they exist in the population. Random sampling of the population ensures that quantitative characterisations of the sample will be generalisable to the population, or at least the population included in the sampling frame. If the representativeness of the sample is in some way compromised (for example, by differential rates of non-response between subgroups), it may be possible to adjust for this at the analysis stage, by adding greater weight to responses from under-represented groups (Fowler 2002, p48).

The concept of theoretical sampling was introduced by Glaser and Strauss (1967) as part of their 'grounded theory' paradigm of social research. In theoretical sampling, each decision as to whether to sample a particular case remains within the control of the researcher, and will be informed by the particular concerns of the research. The overarching aim of theoretical (or purposive) sampling is that the sample should "include the full range of individuals or cases so that the concepts and categories developed provide a comprehensive conceptualisation of the subject" (Boulton and Fitzpatrick 1997, p84). For example, researchers may set out to sample for diversity, aiming to ensure that views are gathered from all in a list of groups specified in advance; sampling may aim for theoretical saturation, and continue up to the point at which no new insights are available from the data; or it may aim for extreme cases or cases which disconfirm prior or emergent hypotheses (Onwuegbuzie and Leech 2007). Theoretically derived samples will generally not support quantitative forms of generalisation, but they may be intended to facilitate conceptual or interpretive generalisation, by attempting to include all (or an adequate range) of the perspectives within the population.

The implications for diversity of theoretical sampling methods are hard to define. Unlike in statistical sampling, diversity in theoretical sampling is inappropriately measured quantitatively (at least at the level of individual studies), since the goal is not to obtain a sample which is as diverse as the population, but one which is sufficiently diverse to obtain an adequate understanding of the subject under study (Allmark 2004). This is not to say that questions of diversity are irrelevant to theoretical sampling methods. Such methods have the potential to promote diversity, for example, by allowing researchers to focus on specific minority groups who might be marginalised if sampled in proportion to their numerical weight. In general, the key question – in terms of both diversity and best research practice – is whether researchers include a sufficient range of possible cases adequately to support the analyses carried out.

Questions relating to sampling strategies for intervention studies seem to have received relatively limited attention compared with the literature on sampling in qualitative and observational research, although issues relating to the appropriateness of inclusion and exclusion criteria used in trials are now more widely discussed (see e.g. Bartlett et al. 2005, Oakley 2006). A main methodological focus has been on sample size, which has been intensively investigated on account of its connection to statistical power. The relative lack of attention to sampling for trials may reflect a view that who is included in the

sample is a question of external, rather than internal, validity. This is in contrast to survey research, where external validity, and hence the extent to which the sample represents the broader population, is regarded as directly relevant to the central aims of the study.

The sampling strategies employed by the 174 studies in this dataset of HP and PH research are summarised in Table 4.1. We distinguish between sampling of sites or settings (such as schools or social services departments) and that of individual participants. We also make a distinction in the table between four types of sampling: theoretical or purposive; comprehensive; random; and unclear or convenience. Theoretical or purposive sampling is used to obtain participants with particular characteristics, to obtain diverse participants, or to obtain participants who are typical or representative of a broader population. With comprehensive sampling, all relevant settings in a given geographical area, or all participants meeting implicit or explicit inclusion criteria in selected settings, are sampled. Random sampling may be simple or use a stratified approach. In 'unclear' or 'convenience' sampling, researchers either do not say how samples were found or they report using samples that were easy to find.

Table 4.1: Sampling strategies in the views studies (N = 121) and the intervention studies (N = 53)

Sites	Individuals	Views N (%)	Intvn N (%)
Theoretical / purposive	Theoretical / purposive	12 (10%)	0 (0%)
Theoretical / purposive	Unclear / convenience	13 (11%)	3 (6%)
Theoretical / purposive	Comprehensive	5 (4%)	8 (15%)
Theoretical / purposive	Simple random	2 (2%)	0 (0%)
Unclear / convenience	Theoretical / purposive	16 (13%)	1 (2%)
Unclear / convenience	Comprehensive	6 (5%)	18 (34%)
Unclear / convenience	Simple random	3 (3%)	2 (4%)
Comprehensive	Theoretical / purposive	1 (<1%)	0 (0%)
Comprehensive	Unclear / convenience	1 (<1%)	1 (2%)
Comprehensive	Comprehensive	1 (<1%)	1 (2%)
Comprehensive	Simple random	0 (0%)	1 (2%)
Simple random	Unclear / convenience	2 (2%)	0 (0%)
Simple random	Stratified random	1 (<1%)	0 (0%)
Simple random	Comprehensive	1 (<1%)	2 (4%)
Simple random sample from population database		1 (<1%)	0 (0%)
Stratified random sample from population database		1 (<1%)	0 (0%)
Quota-based random sample on street		4 (3%)	0 (0%)
Unclear or purely convenience sample		51 (42%)	16 (30%)

Some examples of these different approaches from the views studies in this dataset are given below:

'The sample was generated from lists of all women who had had a first baby when aged between 16 and 19 in hospitals in the three areas: Hackney, Leeds and Solihull...The records of all 16-19 year old mothers over a period of three to six months in 1995 were identified by hospital staff and scrutinised to ensure that the baby had not died or been adopted and that it was a first baby...This constituted a complete series of first births to 16-19 year old mothers and no sampling took place' (Allen et al. 1998, p 211). This is an example of unclear sampling of sites and comprehensive sampling of individuals.

'The schools were selected to represent a range of geographical areas as well as catchment areas of differing socioeconomic characteristics.' For the questionnaire stage of the study, the sample was all children in relevant year group; for the interview stage, children were selected 'randomly' (Cahill et al. 1996, p 5). This is an example of purposive sampling of sites (for diversity) and comprehensive and random sampling of individuals.

'A sample of ten educational institutions was chosen as being representative of the school system [informally]...Each institution was asked to ensure participation of subjects with low, medium and high academic achievement.' (Gallagher et al.

1992, p 277). This is an example of purposive sampling of sites and purposive sampling of individuals.

'The samples were taken from secondary schools and youth clubs. Both the schools and youth clubs used were situated in varying localities within West Berkshire and Tameside, to access a range of environmental and socio-economic influences' (Lee 1993, p 13). This is an example of purposive sampling of sites and unclear sampling of individuals.

In Table 4.1, the concept of theoretical or purposive sampling has been interpreted broadly, to include any case in which researchers reported an intention to obtain, for example, a range of participants on a particular variable, or a sample which was in some way representative. In many cases, however, no formal strategies were employed to attain this goal, and no information reported which would enable readers to judge whether the goal had been adequately attained. The relevance to diversity of many of these studies is compromised by inadequate reporting of how sampling intentions translated into practice.

Perhaps the most striking feature of Table 4.1 is the large number of studies (N=67, 39%, bottom row of the table) which were either unclear on the matter of sampling or which reported using a convenience sample. Many of the intervention studies tabulated as using comprehensive sampling for individuals could also be included here, as they suggested that selected school classes or year groups were comprehensively sampled without giving explicit details. In general, little detail was available on sampling in the intervention studies. While most indicated who might be eligible to be included in the sample, and some reported inclusion criteria either for the intervention or for the study, none gave a detailed characterisation of the whole process.

One example is an RCT of a primary-school-based intervention to reduce risk factors for obesity (Sahota et al. 2001). Sahota and colleagues report: 'We recruited 10 primary schools in Leeds...All the participating schools were state primary schools sited outside the inner city area. Socio-demographic measures suggested that the schools' populations generally reflected the Leeds school aged population, although there was a slight bias towards more advantaged children. The schools had 1-42% children from ethnic minorities and 7-29% entitled to free school meals compared with 11% and 25% respectively for Leeds children as a whole' (Sahota et al. 2001, pp 1029-1030). While the description of baseline demographics here increases the study's external validity, it remains unclear why the participating schools were chosen for inclusion in the study rather than any others. The authors do not report whether any schools declined to be included, nor whether the exclusion of inner-city schools was a deliberate sampling decision. The lack of detail here contrasts with the detailed attention paid in the report to methodological issues relevant to internal validity, such as randomisation and attrition. Studies of the implementation of research in the HP and PH field have emphasised the need for clearer reporting of sampling decisions and results at the level of both individual participants and institutional settings, such as schools or health care units (Estabrooks et al. 2003, Glasgow et al. 2002).

Another interesting result is the number of studies (N=33, 19%, rows 2 and 5 of the table) which reported either purposive sampling of sites and convenience sampling of individuals, or *vice versa*. Where the sites sampled from which samples are taken are not diverse, sampling participants who are diverse relative to these sites is of limited use, and conversely, many of the benefits of sampling sites for diversity are likely to be lost where barriers to diversity at the individual level are not addressed. Strategies such as deliberate over-sampling of

disadvantaged or minority groups, as reported by four studies in our dataset (Armstrong et al. 1998, Bostock 2001, Chinn et al. 2004, Honess et al. 2000), may be valuable in addressing this issue.

Many of the studies claimed that either the sites or the individuals sampled included diverse groups, but reported no details of who was included, or of the sampling procedure. For example, one study of young people's views of health and exercise (Harris 1994) reports that: 'Sixty-one young people (11-14 years) from two different state secondary schools ... were involved in 14 focus groups. The schools were both large comprehensive schools, one in a main town and the other in a small town lying on the outskirts of an expanding major town. Both schools attracted pupils from urban and rural settings and from a mix of socio-economic backgrounds' (p 144). Such reports evince a commitment to diversity as a valuable feature of research, without this appearing to inform the sampling strategy. This loose relation between intention and method is frequently encountered in this dataset. In total, 57 of the 121 views studies (47%), as well as 5 of the 53 intervention studies (9%), stated some intention to obtain a sample which was either diverse or representative of some broader population, but relatively few made clear how they intended to, or did, achieve this.

Two questions may arise with any form of sampling method. The first concerns whether the diversity agenda is better served by comparing diverse groups within a population, or by focusing attention on particular groups without making any such comparisons. Avoiding inter-group comparisons, and concentrating on sampling one particular group, may assist in gaining an 'emic' or 'insider' perspective on the group under study. However, sampling only one group does not allow researchers to contextualise their findings by distinguishing what is shared between groups and what is particular to a given group.

The second question is what constitutes an adequate sample size. In general, sample size should be driven by the type of analysis the data is intended to support. Where data is to be tested statistically, power calculations will show what sample size is required for particular tests; the sample size should also be sufficient to power any subgroup analyses or interaction tests which are to be carried out (Brookes et al. 2001). Where data is to be analysed qualitatively, formal power calculations cannot be used, although some writers have outlined analogous methods (Onwuegbuzie and Leech 2007). Atkin and Chattoo (2006 p104), considering this question from the perspective of qualitative research on ethnicity and health, write: 'As a rule of thumb, sample size must address both the diversity within an ethnic group and yet make theoretical generalization for subgroups meaningful. However, there is no statistical logic of validity informing the right number of individuals/cases, and these issues are addressed theoretically in relation to the kind of data required as well as approach to analysis'. In other words, considerations of sample size (how many cases should be sampled) overlap with the considerations of sampling methodology outlined above (which cases should be sampled): the sample must include both *enough* cases and *the right* cases to support inferences about both how diverse groups differ and what they share.

4.3.4 Sampling frames

Interestingly, the choice of a sampling frame seems to have received less attention in the theoretical literature from the point of view of potential for bias than has the choice of method to sample from the frame; the empirical literature

on the effect of different sampling frames is limited (Wolf et al. 2005). Very little work has been done specifically focusing on children and young people.

Statistical sampling methods, such as random sampling, if carried out rigorously, are effective in ensuring that the research sample adequately reflects the diversity to be found in the sampling frame. However – apart from threats to the diversity of the sample resulting from differences in recruitment, consent and attrition rates between different groups, or by the failure to recruit particular groups, such as non-English-speakers – the sampling frame may also itself be unrepresentative of the population under study. Researchers may sample only from a limited geographical area, or from a limited range of particular settings (e.g. schools). Sampling strategies which aim to cover the population as a whole must often rely on frames, such as electoral registers or telephone directories. Many of these information sources are problematic, in that they under-represent ethnic minorities, lower-income people, and other marginalised groups (Sin 2004). While adjustments can be made so as to oversample minority groups (Kalsbeek 2003), the difficulty of anticipating the effect of all relevant dimensions of diversity means that not all the potential limitations resulting from the choice of a sampling frame can be overcome.

Some statistical sampling strategies do not require a sampling frame, and hence are not subject to some of these limitations (Fowler 2002). For example, households can be recruited door-to-door, using area probability sampling or similar methods to obtain samples which are representative of the population; alternatively, random-digit telephone dialling can be used, although telephone samples are generally not as representative (Brogan et al. 2001). Such methods are probably the most effective way of obtaining samples which are representative of the general population, and studies using these methods have obtained diverse samples of disadvantaged populations of children and young people (e.g. Siegel et al. 1999). However, non-response rates, while generally lower than those achieved with other methods, are not reduced to zero, and may still affect the diversity of the sample. Both methods involve certain exclusions: household surveys omit homeless people and those living in institutions such as prisons, and telephone surveys exclude people who do not own telephones.

4.3.5 Sample recruitment

A considerable body of literature exists on the best ways to recruit minority populations, particularly ethnic minorities, into research (for reviews of this literature, see Hussain-Gambles et al. 2004, Swanson and Ward 1995, Yancey et al. 2006). Most of this research has focused on adult populations. A central issue is the settings used to recruit participants. Table 4.2 shows the settings from which research participants were recruited in the sample of 174 HP and PH studies relating to children and young people.

Table 4.2: Settings used to recruit participants (N = 174 studies; some studies used more than one setting)

	N (%)
Via schools	109 (63%)
Via social services / welfare / family unit	21 (12%)
Via health services / clinician referral	15 (9%)
Community site / outreach	9 (5%)
Youth clubs / sports clubs	8 (5%)
Via community organisation	8 (5%)
Personal contact / word of mouth	5 (3%)
'Snowballing'	5 (3%)
Mass media	5 (3%)
Prison / youth offender teams	4 (2%)
Population registers	4 (2%)
Workplaces	3 (2%)
Children's homes	1 (<1%)
Unclear	21 (12%)

Twenty-one (12%) of the studies were unclear as to the settings from which the samples were recruited. The majority of studies (N=109, 63%), and the overwhelming majority of those which sought a general-population sample of school-aged children or young people, recruited wholly or primarily through schools or colleges, although only one of these clearly reported a rationale for schools rather than other settings being chosen (Bendelow et al. 1998, p 5). The availability of schools as settings for research means that many of the barriers to diversity encountered when recruiting adult participants to research are less pressing for researchers working with children and young people. However, certain barriers to diversity with school-based recruitment remain. It seems likely that schools whose staff are enthusiastic about participating in research are included in research much more than those whose staff are not, and this enthusiasm may be correlated with the socio-demographic profile of the school. However, the lack of detailed description in research reports (of trials as well as views studies) as to how schools were sampled makes it difficult to quantify the bias arising from school selection.

Diversity may also be compromised because not all children of school age attend school, or will be attending school on a particular day. School exclusions and absences affect certain groups disproportionately, with absences showing a particular gradient for SEP (National Statistics 2007a), and exclusions for gender and ethnicity (National Statistics 2007b). Some children are also educated outside school; several potentially marginalised groups – such as travellers, children with long-term illness or with a family member with long-term illness, and possibly members of certain cultural or religious groups – are more likely not to attend school or to attend for only some of the year (Scottish Executive 2006).

One approach to this issue is to supplement school-based recruitment with 'booster samples' from other sources, such as community groups, specialist services (health or social services) and youth clubs. This method has been widely used in population-based quantitative surveys (such as the British Crime Survey), and in some partly school-based studies of social exclusion and substance use, such as the Belfast Youth Development Study (McCrystal et al. 2007). Only one study in this dataset adopted such an approach (Armstrong et al. 1998). One potential barrier to diversity where participants are recruited either through schools or through other services is the role of 'gatekeepers' – such as teachers or other service providers – in recruitment. Such gatekeepers may be influenced in whom they recruit by conscious or unconscious prejudices regarding who is suitable for the research: for example, they may exclude children or young people judged to be shy or potentially disruptive, or those (in an intervention study) considered by them less likely to benefit from the intervention. In some cases, gatekeepers may restrict access from a desire to protect those in their care (Atkinson and Flint 2001). A few studies in this dataset reported explicitly that gatekeepers were forbidden by the researchers to 'pre-screen' participants, and required to recruit eligible participants wherever possible (Hahn et al. 1994). One study reported that the recruitment practices of gatekeepers (in this case, health visitors) biased the sample towards lower-SEP participants (Speak 1995).

Recruitment methods based on children or young people self-selecting or volunteering to take part in research may produce a sample which is unrepresentative of the sampling frame as a whole. Of the school-based studies in this dataset, seven (6%) explicitly stated that participants actively volunteered to take part. It is often difficult to tell from research reports whether participants were volunteers, were selected by teachers, or included all children – in particular, classes who were attending school on the day; in a few cases, where both recruitment and data collection were carried out by school staff, researchers may themselves have been uncertain.

Studies which recruit from schools often report demographic characteristics of the school but not of the actual sample. For example, 72 (60%) of the 121 views studies in this dataset stated that the sample was obtained wholly or partly from schools, and a further seven (6%) implied that the study was school-based without giving information on sampling methods. Of these 79 studies, 39 (49%) gave some information on SEP, but only 16 of these (20%) reported on the SEP of the sample itself; the others described only the SEP of the school, or of the community in which it was located.

Where research which is specifically targeted towards hard-to-reach or socially excluded populations, school-based recruitment will often be inappropriate. Some populations (such as young mothers) can be accessed through other statutory organisations, such as health services or social services, or voluntary or community organisations. Most of the non-school-based studies in this dataset used this type of approach. 'Gatekeeper' effects may be particularly pronounced with this approach, since issues of trust – between participants and researchers, and between participants and institutional service providers – are particularly salient with socially excluded populations (Emmel et al. 2007). Additionally, in some cases, non-users of services may differ systematically from users. Other populations (such as refugees and asylum-seekers, or drug users) may be largely 'invisible' to statutory or voluntary services and require highly targeted recruitment strategies. Here strategies such as 'snowballing' – where existing participants recruit friends or acquaintances – may be valuable. However, these methods are highly susceptible to bias (Faugier and Sargeant 1997). It may be useful to

implement quotas for particular groups, although such methods are not statistically rigorous (Bloch 2007). More sophisticated approaches, such as respondent-driven sampling, allow the integration of probability sampling with snowball-type methods, helping to overcome the potential for bias (Magnani et al. 2005).

Related to recruitment are issues of retention and drop-out, particularly in longitudinal studies. It is known that adult participants from disadvantaged groups are often more likely to be lost to follow-up (De Graaf et al. 2000, Hille et al. 2005, Siddiqui et al. 1996). Researchers conducting longitudinal studies should be aware that even where initial samples reflect the diversity of the population, differential attrition may compromise the diversity of the sample. We did not find any discussion of this issue in the methodological literature in relation to children and young people in particular. It is possible that drop-out is less likely when samples are school-based, compared with community-based, but this issue requires further investigation.

4.3.6 Consent

Bias towards unrepresentative samples may also occur as a result of participants, or their parents, refusing consent (Klepp 1995). Using passive rather than active consent procedures – allowing participants to opt out rather than requiring them to opt in – may increase recruitment of ethnic minorities and other disadvantaged groups (Testa and Coleman 2006, Unger et al. 2004), although active consent is often considered ethically preferable (Alderson and Morrow 2004), and in many cases may be required by funders or governance procedures.

The consent procedures used may themselves pose barriers to diversity. This is another aspect of research where poor reporting procedures make it difficult to judge whether consent was sought and, if so, how the consent process might have encouraged or impeded the goal of diversity. For example, of 53 studies in this intervention dataset, 24 (45%) did not clearly report that consent was sought. Recent research has identified the risk that overly restrictive conceptions of 'informed' or 'competent' consent may unjustly exclude from research individuals or groups who are in fact capable of understanding the implications of their decision, including young children (Alderson and Morrow 2004, Cocks 2006, Morrow and Richards 1996), children with learning disabilities (Cocks 2006) and people with mental health problems (Osborn 1999). Research on the socio-cultural determinants of research participation has emphasised that the role of ethical safeguards in protecting potentially vulnerable groups should not imply the exclusion of minority groups who are presumed to be unlikely to consent to research (Ashcroft et al. 1997, Levine 1991).

4.3.7 Data collection

As with other stages of the research process, the question of how best to collect data from diverse populations overlaps with the question of what is best practice in research more generally. Data collection should have the aim of allowing all participants to contribute equally to the research findings. This means that, for example, data collection tools need to be appropriate for all groups in the population under study. Written questionnaires or other methods dependent on literacy in English pose an obvious problem for non-English speakers, or for those with English as a second language. Outcome measures may embody culturally loaded assumptions. Even widely-used standardised tests may not have been

validated with all populations, and they may not capture the outcomes which are most relevant to participants themselves.

In this intervention studies dataset, 30 out of the 53 studies (57%) stated that the tools used to measure outcomes had been previously validated, although only one explicitly said that such validation had been carried out with diverse populations (Gortmaker et al. 1999). Five studies stated that data collection tools were piloted with a sample of the population before being used in the full study (Liquori et al. 1998, Luepker et al. 1996, Moore 2001, Nicklas et al. 1998, Perry et al. 1998). Of the 121 views studies, 28 (23%) reported that some piloting or pre-testing of data collection tools had been carried out, although again, only one explicitly said that validation was with diverse groups (Armstrong et al. 1998).

Three dimensions of diversity are particularly relevant here: ethnicity (insofar as it overlaps with culture and language), age, and disability (where this includes learning disability and/or communication impairments). The question of what methods are appropriate for younger children, or for groups of mixed age and/or ability, has been widely addressed in the literature (Hatch 1995, MacNaughton et al. 2001). Partially non-verbal methods, such as games or drawing, are valuable with younger children. For children and young people with complex communication impairments, specialist tools (such as computer-assisted interview packages) are available (Knight et al. 2006). In this dataset, two intervention studies (Anderson et al. 2003, Perry et al. 1987) and six views studies (Bickerstaff and Shaw 2000, Birtwhistle and Brodie 1991, Mauthner et al. 1993, Mulvihill et al. 2000, Taylor and Halliday 1996, Tyrrell 2000) offered detailed rationales for their data collection methods in terms of what was appropriate for the age group under study. Two studies (Bickerstaff and Shaw 2000, Tyrrell 2000) reported extensive use of multiple methods to ensure that older and younger participants, and participants of mixed ability levels, could contribute to the research. Two studies where the primary data collection tool was a written questionnaire reported that researchers confirmed that participants had understood the questionnaire (Bowen 1997, Gibson et al. 1998). No studies reported that data collection methods were chosen or adapted for disabled or communication-impaired participants.

These broad issues of the accessibility and appropriateness of tools apply to the collection of any type of data. Two more specific issues in research using interviews or focus groups concern the role of diversity within the research situation. One debate is whether the 'matching' of interviewer and interviewee(s) – that is, their being the same gender, ethnic group etc. – provides more reliable data. The literature on so-called 'race-of-interviewer effects' is now considerable, and seems to indicate that, especially on sensitive topics, ethnically matched interviewers do elicit more detailed data (for an overview, see Grewal and Ritchie 2006). However, these results cannot be applied mechanically to a given interview situation, not least on account of the complexity of individuals' ethnic and other identifications (Flaskerud and Nyamathi 2000, Gunaratnam 2003). In some situations, the non-matching of interviewer and interviewee may have a disinhibiting as much as an inhibiting effect (Grewal and Ritchie 2006, Phoenix 1994). The researcher's own ethnicity (or other dimension of diversity) appears to be less important than her/his cultural competence and sensitivity (Flaskerud and Nyamathi 2000, Sawyer et al. 1995). In this dataset, only one study stated that interviewers and participants were matched for ethnicity (Armstrong et al. 1998), and in two studies there was explicit matching for gender (Denscombe 1999, Honess et al. 2000).

The second debate is whether, in group interviews, researchers should separate diverse groups, for example interviewing males and females separately, or interviewing ethnic minority young people separately from others. In this dataset, 14 studies which used group interviews separated participants: 12 studies stated that group interviews were carried out at least partially with single-gender groups and in six studies groups were matched for other dimensions of diversity, such as ethnicity or SEP (Armstrong et al. 1998, Danton et al. 2003, Engineer et al. 2003, Hughes et al. 1999, Jones 1998, Scott Porter Research and Marketing 2000). However, few of these studies provided detailed rationales for separating diverse groups at the data collection stage. While in certain types of research it may be necessary to separate participants by characteristic to avoid the risk of harm, this does not appear to have been the case with any studies in this dataset.

A data collection method which may be of value in eliciting data from members of marginalised groups is to carry out group interviews of participants in friendship groups; this method was adopted by several studies in this dataset (Barrell 2000, Mitchell 1997, Scott Porter Research and Marketing 2000). A step beyond this is to work with peer interviewers both to assist with recruitment and carry out data collection (Hall et al. 2003); however, such methods pose both scientific and ethical challenges to researchers (Elliott et al. 2002).

4.3.8 Data analysis

Data analysis in HP and PH research can take a number of forms. In this section, the focus is primarily on statistical significance testing and thematic analysis.

Where data is analysed using statistical tests, the principal question relating to diversity is whether subgroup analyses of outcomes or variables of interest should be carried out according to socio-demographic variables. Some researchers argue that carrying out such subgroup analyses is an important tool for monitoring diversity and inequalities. In the USA, the NIH guidelines on inclusion of women and minorities and research state that 'inclusion of the results of sex/gender, race/ethnicity and relevant subpopulations analyses is strongly encouraged in all publication submissions'. Such analyses should be carried out even when 'the data from prior studies strongly support no significant differences of clinical or public health importance in intervention effect based on sex/gender, racial/ethnic and/or relevant subpopulation comparisons' (National Institutes of Health 2001). This position holds that no specific rationale is required to analyse data in terms of diverse population subgroups, and such analyses should be a matter of routine. Another position in this debate holds, by contrast, that subgroup analyses should take the form of hypotheses specified in advance, and for which some rationale can be given (Oakley 2006). Analysing data by subgroup without an adequate rationale amounts to data dredging and is likely to generate misleading results (Brookes et al. 2001, Rothwell 2005b). The routinisation of subgroup analysis risks reifying an arbitrarily chosen set of inter-group differences (e.g. between genders or ethnic groups) which may in many cases have no impact on outcomes; this practice may constitute a form of hidden discrimination.

In this dataset of 53 HP and PH intervention studies, 27 studies (51%) provided some analysis of outcomes by subgroup, or used subgroup variables as covariates in statistical analysis (see Table 4.3). Studies varied in the extent to which these subgroup outcome data were reported. Some studies (N=7) either reported only significant subgroup differences, or stated that no significant differences had been found, and did not provide substantial subgroup data. It was

often unclear to what extent subgroup analyses were driven by pre-specified hypotheses. Several studies offered a rationale for subgroup analysis in terms of the external validity of the study or the generalisability of the intervention. However, in many of the studies which did provide subgroup outcome data, no explicit rationale was given, and none of the studies mentioned subgroup analysis as a primary research aim.

Table 4.3: Dimensions of diversity in analysis of outcome data in the intervention studies using subgroup analysis (N = 27; some studies used more than one dimension)

	N (%)
Gender	20 (74%)
Ethnicity	8 (30%)
SEP	6 (22%)
Educational level (of participants)	4 (15%)
Family / marital status	3 (11%)
Health status (e.g. obese vs. non-obese)	3 (11%)

Gender was the main dimension of diversity used in these subgroup analyses; it was used more than twice as often as ethnicity or SEP, although frequently without any scientific rationale being offered.

A further concern with some statistical analyses is when socio-demographic variables relevant to diversity are used only as confounders in the analysis to control for their effect on the outcome variable of interest. While such analyses are valuable, focusing on 'adjusting out' the effects of diversity may represent a missed opportunity to deepen an understanding of the processes which lead to differences in outcomes between diverse groups.

In the case of qualitative research, the question of how data analysis can best promote diversity is less clear-cut. The basic issue, however, is the same: analysis should explain patterns of variation in the data, without introducing artefacts or imposing inappropriate structures upon them, and it should produce findings which are grounded in the experiences of all participants. Some quality assessment tools for qualitative research ask reviewers to assess how well diversity in perspective has been explored (Long and Godfrey 2004, Spencer et al., 2003). As with statistical analysis, thematic analysis needs to be tied securely to the data. A common feature of analysis and reporting in qualitative research is the use of direct quotations from interviews/focus groups. In the dataset of 121 views studies, 78 reported such direct quotations from participants. However, only 43 of these (55%) consistently reported some information on the socio-demographic characteristics of the individuals quoted. In most cases, this was limited to age (N=23, 29%) and/or gender (N=28, 36%); only four studies (5%) reported qualitative data by ethnicity, and five (6%) by SEP or occupational status. Of 86 views studies which included some quantitative data (such as percentages), 23 (27%) reported all such data by gender, 15 (17%) by age, four (5%) by SEP and three (3%) by ethnicity. A small number of further studies provided incomplete socio-demographic information relating to quotations (N=7) or quantitative data (N=27), but did not report all data in this way. In other words,

relatively little information can be gained from research reports regarding what views were expressed by whom.

The question as to what justifies the choice of particular variables to use in the analysis is important in qualitative as well as quantitative analysis, since the open-ended nature of interpretive analytic methods creates considerable scope for researcher-imposed artefacts in the findings inferred from the data. In this views dataset, almost half the studies (N=58, 48%) gave no clear description of the method of data analysis used. Many of those which described the data analysis process provided only broad descriptions. Relatively few studies explicitly reported methods used to promote diversity at the data analysis stage. One study, which included a specific sample of ethnic minority young people, stated that the findings of the analysis were checked with the (ethnically matched) community workers who carried out the interviews (Armstrong et al., 1998). Another study reported that quotations were selected for inclusion in the report in proportion to the numbers of participants making similar comments (Allen et al. 1998). Five studies said explicitly that dimensions of diversity such as age, gender and/or ethnicity were taken into account in the process of interpretive analysis (Coakley and White 1992, Coleman and Cater 2005, Hart et al. 2002, Honess et al. 2000, Jones, 1998). A further two studies reported such dimensions emerging as explanatory variables in the course of analysis, without having been posited at the outset (Davis and Jones 1996, Harris 1994). Many more studies reported data according to dimensions of diversity, and/or mentioned particular dimensions of diversity in the presentation of their findings, without explicitly including these as factors in the analysis. Most reports which described data analysis made some commitment to the principle to analyse data so as to represent participants' own views rather than impose other interpretations. However, there were few guarantees in these reports that this principle was respected in the analysis which was actually carried out.

4.4 Conclusion

Researchers' concern for diversity or representativeness as goals of HP and PH research often does not carry over into research practice. In many respects, the promotion of diversity follows from good practice in research more generally. Existing guidance on the conduct of research already constitutes a powerful framework for the development of research practice, even if this perspective has rarely been explicitly addressed in published guidelines or quality assessment tools. The promotion of diversity requires an ongoing commitment by researchers to critical reflection on their own practice.

5. DISCUSSION AND RECOMMENDATIONS

The review described in this report aims to answer the question as to how health promotion (HP) and public health (PH) research relating to children and young people has addressed issues of population diversity. The rationale for inclusive research procedures is both scientific and ethical. Research findings drawn in appropriate ways from diverse samples have greater generalisability and relevance to policy, and they also reflect an ethical concern for people's rights to participate in research whose findings may affect their welfare and wellbeing.

In this report, we have discussed the rationales for including diverse groups in research, and considered what strategies for inclusion might be appropriate for different groups, different research questions, and different stages of the research process. The approach developed to examine the literature was based on two frameworks: first, an open-ended list of dimensions of diversity based on the six statutory strands of equality included in the Department of Health Single Equality Scheme (Department of Health 2007), but also including socioeconomic position; and second, a schematic outline of different stages of the research process. We also analysed a specific dataset of studies included in previous systematic reviews of HP and PH research relating to children and young people.

5.1 Principal findings

Our main findings are that poor reporting hinders judgements about the extent to which research in this area reflects population diversity. More than two in five of the 174 studies examined gave no information about SEP, and more than half provided no guide as to participants' ethnicity. Information on gender, age or number of research participants was missing in around one in 20 of the studies. Comparing the two types of studies in this dataset – trials of HP and PH interventions and studies of children and young people's views and experiences – it was clear that the intervention research did a better job of reporting than the views studies. For example, 96% of the intervention studies, compared with 88% of the views studies, provided exact numbers of research participants; and 68% of the interventions studies, compared with 18% of the views studies, gave precise information on ethnicity.

There are different stages of the research process where an inclusive or non-inclusive approach is relevant. It was sometimes hard to tell from researchers' descriptions how samples had been located, recruited or informed about the study, and whether or not diversity had been foregrounded in the choice of data collection tools, measures or data analysis methods used. For example, 67 studies (39%) did not report any formal sampling strategies, and many of the rest offered only vague and incomplete descriptions. A common approach was to claim that research samples were socially diverse, but provide no details. Unsurprisingly, given the focus of the 174 studies on children and young people, schools were the main setting used for sample recruitment, but information was

often missing on which schools had been chosen and why, and frequently socio-demographic data was provided only at the school and not at the individual level. Around half the intervention studies (N=24, 45%) did not say anything about consent procedures. Only two studies in the dataset explicitly reported the use of tools validated with diverse groups. A common data analysis practice in qualitative research is to use direct quotations from interview or focus group material. Seventy-eight of the 121 views studies adopted this approach, but only 43 (55%) of the 78 consistently attached any socio-demographic descriptors to the quotations.

The extent to which diversity is reflected in HP and PH research is, and must be, partly shaped by the questions addressed in particular studies. In an intervention trial focused on strategies for helping obese children, for example, those social groups with the highest risk of obesity will be over-represented, and the study sample will hence not reflect the social and ethnic mix of the community. A views study which asks about young people's experiences of teenage pregnancy will under-represent those young women from minority ethnic groups which have low rates of teenage pregnancy. Forty-nine of the 174 studies (28%) in this dataset targeted specific populations, a strategy which necessarily puts limits on the extent to which these studies can be expected to include socially representative samples.

While research targeted at disadvantaged groups has an important role to play in the promotion of diversity, there are risks associated with over-emphasising such an approach. One issue is that it may contribute to a culture in which disadvantaged groups are seen in terms of the problems they pose, and such groups may view research itself as potentially stigmatising (Chase et al. 2003). There is also a danger that this approach may lead to a weakening of the inclusion and diversity agenda overall: where the burden of including diverse groups is effectively shifted to targeted research, there is a risk that the findings of research concerning the population as a whole will have less incentive to take into account either the exclusion of particular groups from their samples, or the internal diversity of those samples.

A key issue is the extent to which researchers offer a convincing rationale for the strategies they have used to include or exclude a focus on diversity. For example, sample recruitment in most studies in the UK and the USA is restricted to English-speakers, a strategy which discriminates against all would-be participants who lack fluency or literacy in English. However, very few studies note this restriction, and even fewer comment on the implications it may have for the character of the sample and the nature and generalisability of the research findings. Another area in which rationales are thin on the ground is in relation to subgroup analysis. This is a common feature of intervention studies: 27 (51%) of the 53 studies in this dataset used it. Despite established guidance on good practice with respect to subgroup analyses in trials (Brookes et al. 2001, Rothwell 2005b), few of these studies gave a convincing explanation as to why they undertook particular subgroup analyses.

Our findings parallel those of other studies which have looked at how different types of studies are designed and reported: see, for example, Bartlett et al. (2003), Clay et al. (2002), Sheikh et al. (2004), Silagy and

Jewell (1994), Waldenström and Turnbull (1998). For example, with respect to consent, the findings showed that 45% of the intervention studies did not clearly report that consent was sought; in the earlier analysis of 215 HP and PH intervention studies, this figure was 73% (Oakley et al. 1998). A review of 489 school-based studies covering both education and HP and PH research found that 70% either did not say anything about consent or reporting was unclear (Rees et al. 2007). In an earlier analysis of trials published in the *Archives of Disease in Childhood*, Campbell and colleagues (1997) reported that 45% gave no information about whether or not consent had been sought. In the analysis of educational research by Newman and Elbourne (2004), 75% studies did not report whether or not consent had been sought. Other findings from this analysis echo ours: 8% of studies lacked precise information about the number of participants; the ethnicity and SEP of individual study participants were not clearly stated in 67% of the studies; information was missing or incomplete on the gender of participants for 19% and on age for 7%. Other researchers report similar findings to ours on the lack of information about sampling and recruitment strategies (Gross et al. 2002, Newman and Elbourne 2004).

5.2. Frameworks for good practice in research

As noted above, the inconsistent and patchy picture we found with respect to the coverage of diversity issues in this group of HP and PH studies contrasts with a context in which there is now a substantial body of literature providing structured guidelines for good practice in the design, conduct and reporting of various types of research. For trials there are the CONSORT guidelines (Moher et al. 2001), with a version for cluster trials (Campbell et al. 2004) (which pose complex issues with respect particularly to consent and data analysis). However, neither of these frameworks go further than requiring the reporting of eligibility criteria. As Dieppe and colleagues (2004) note in their analysis of trials of commonly used drugs for osteoarthritis, there has been much less attention in trials to the problem of external, than to that of internal, validity. Guidelines have been proposed for educational research (Newman and Elbourne 2004) which include specifications for the provision of detailed socio-demographic information relating to different stages of the research process. Although, as noted in Chapter 4, there is a debate about whether the validity of qualitative research should be judged in the same way as that of quantitative/experimental research, there is a burgeoning literature on checklists and criteria for judging the reliability and representativeness of this type of research: see, for example, Bornhöft et al. (2006), Badger and Werrett (2005), Dixon-Woods et al. (2004), Miller et al. (2003), Parkes et al. (2006), and Spencer et al. (2003). A recent systematic review of checklists for the reporting of qualitative research found 22 such checklists (Tong et al. 2007). Journal editors also have a role to play in ensuring good reporting of research, but few social science journals have established guidelines for good reporting practice.

There is a growing commitment in the research community to improve the conduct and reporting of research. This includes the need to match research samples more closely with the population to which research findings may be generalised. Information about the match or mismatch

between research samples and the populations they supposedly represent often requires extensive detective work. For example, a study by Mason and colleagues (2003) of ethnic representation in trials used participants' names to construct an ethnic profile; this showed that the presence of South Asian people in the trials under-represented by some 50% their prevalence in the population.

The comparison in Chapter 4 of the methodological framework for good practice with the findings from our empirical analysis of how HP and PH research is actually done reveals considerable discrepancies. For example, while the literature on theoretical sampling methods includes a number of strategies which are valuable from a diversity perspective, very few studies which used suitable designs clearly report how sampling decisions were informed (beyond the standard caveat that sampling did not aim for statistical representativeness).

The overall picture which emerges from this discussion of the methods literature is that the inclusion of diverse groups and the conduct of rigorous research go hand in hand. Specific recommendations which go beyond this – such as the recruitment of very hard-to-reach populations, or ensuring that all resources can be understood by all minority groups – may be highly important in particular research fields, but are likely to hinder, more than help, in general-population research, where 'the excellent may ... be the enemy of the good' (National Co-ordinating Centre for Research Methodology 2001, p 8).

5.3 Strengths and limitations of this review

The review presented in this report attempts to draw together material relating to diversity in research with children and young people from a range of fields. It is unusual in that it combines a reanalysis of empirical studies with a literature review of work on diversity and research practice. This has the advantage of enabling us to place the findings of their analysis of empirical studies in the broader context of debates among researchers about the relationship of social diversity to research practice.

Both the literature review and reanalysis of studies had their limitations. In both, we relied on studies published in the English language, a constraint that is somewhat paradoxical in the light of the focus of the review. For the reanalysis, we relied on studies from a series of previous reviews. This meant that some studies dated from several years ago, with several (N=36, 21%) published prior to 1995, although the majority were relatively recent, with 138 (79%) published in 1995 or later, and 66 (38%) in 2000 or later. In addition, although each individual review was systematic, and collectively they covered a broad range of key topics in HP and PH research, none (with the exception of the review on teenage pregnancy and parenthood) was specifically focused on diversity or on disadvantaged populations. The alternatives to sourcing studies from previous reviews would have been, first, to undertake systematic database searching for primary research studies; or, second, to identify one or several key journals and include all relevant studies published in that journal(s) within a given timeframe. While both approaches have certain advantages, it is unlikely that either would have facilitated the

inclusion of as broad a sample of studies as were included in this reanalysis. This dataset included a wide range of study types and methodologies, populations and settings, and incorporated general-population research without a diversity focus, rather than focusing exclusively on research targeted at disadvantaged populations.

We included intervention studies judged methodologically sound in the original reviews, but the views studies varied in their methodological quality. Had they included intervention studies assessed as methodologically weaker, their findings might have been different. Indeed, their previous analyses of the socio-demographic characteristics of samples included in EPPI-Centre HP and PH reviews (Oakley et al. 1998, Oakley 2006, p 145) showed a clear relationship in intervention research between quality of reporting and methodological quality.

Our searches for methods literature were far from comprehensive, and were driven by emerging conceptualisations of the field rather than by an *a priori* delimitation of the relevant literature: hence they will certainly not have located all relevant material. The material which was located was heterogeneous: certain themes have been extensively and rigorously researched, while much work remains to be done in relation to others.

This review is paralleled by a second EPPI-Centre review undertaken at the same time which is focused on inequalities in HP and PH related to children and young people (Oliver et al. 2008). The two reviews are complementary: a focus on health inequalities requires attention to issues of social diversity in the planning, design, conduct, analysis and reporting of research; and attention to diversity issues invariably reveals health and social inequalities. Part of the health inequalities review consists of a detailed look at how HP and PH researchers undertaking intervention evaluations have described populations and measures of disadvantage, and have operationalised the assessment of gaps and gradients in health status. The findings of the inequalities review complement those of the diversity review in demonstrating rather more rhetoric than critical and productive reflection about the challenges of designing research capable of making a useful and reliable contribution to the HP and PH evidence base.

The broadening of perspective beyond individual studies places a strong emphasis on the role of research synthesis in understanding the existing evidence base regarding diverse groups. This is one reason why the routine reporting of demographic variables in primary studies is valuable: it enables researchers to use meta-analysis or qualitative synthesis techniques to identify patterns in the data which may not have been clear in individual studies.

5.4 Recommendations

Good practice in research with respect to inequalities and diversity requires both prescription and flexibility. Standardised lists of issues to be addressed in designing and reporting research are helpful, but research questions and designs vary. No single research project can be expected to address all aspects of diversity, and attempting to do so would

frequently be counter-productive: 'the principle is one of inclusion. It must be reconciled with the aim of delivering high quality meaningful outputs and maximising value from research funds. It is unethical, as well as poor value for money, to involve participants in research that has not been designed to achieve a worthwhile outcome' (Department of Health 2003, p15).

Based on our overview of methods literature and their dataset of research reports, we suggest below some questions and points of reflection for researchers, funders and journal editors to follow which could extend awareness of diversity issues and thus broaden the relevance of research to a diversity-sensitive HP and PH policy. Although the main focus in this report has been health promotion and public health for children and young people, the key findings and recommendations apply to research in any area across all age ranges. We first offer some specific recommendations; Table 5.1, which follows, lists broad diversity-related goals against the different stages of research.

5.4.1 General recommendations

Reporting

All researchers should present baseline demographics according to major demographic variables (age, gender, ethnicity, SEP). Such a recommendation need not be motivated by a belief that these variables are always relevant, or that outcome data should always be presented or analysed according to them. Reporting of baseline demographics is of value because the external validity of studies will be strengthened by clear reporting of who was included in the sample. What is known about the health of diverse groups depends on having a strong evidence base of relevant studies.

More extensive reporting of, and reflection on, sampling procedures would be valuable in identifying barriers to diversity in trials, and increasing their external validity.

Researchers working with samples of children and young people recruited from schools should report demographics at the individual level as well as the school level.

Targeting particular groups

Researchers should be explicit about what motivates their focus on specific groups and/or their comparisons between subgroups (existing research literature, ethical or political agendas, or themes emerging from the data), and how this motivates the adoption of particular sampling strategies. They should be aware of the limitations of both methods: that comparative methods may miss important dimensions of individuals' experiences, and data drawn from a particular group cannot support inferences about what is particular to that group relative to any wider population.

Recruitment

Research on certain topics may benefit from extra strategies to include certain groups which would be unnecessary for other topic areas. For example, one study from this dataset on young men and mental health included prisoners, homeless people and young men participating in recreational activities (Aggleton et al. 1995); another on underage drinking included young offenders, early school-leavers and looked-after children (Honest et al. 2000).

Data collection

Data collection tools should be appropriate and accessible for all groups within the population.

In order to identify potential barriers to diversity, researchers should validate all data collection instruments with reference to all groups within the population under study, either by piloting or through a consultation process with participants.

Instruments, such as questionnaires, should ideally be validated with both bilingual and monolingual participants, and translations compared with each other (Bhopal et al. 2004). For more in-depth qualitative research, the most effective solution might be for trained bilingual (and bi-cultural) researchers to collect data (Jones and Allebone 1999); however, such researchers should not be taken as representative of minority cultures as a whole (Temple 2002; Temple 2005).

Methodological research

Further methodological work is required on how initial sampling decisions are made in trials, and what methods might be appropriate in different settings.

More work is needed to bridge gaps between theory and practice, and to determine the effect on diversity of commonly used, but under-theorised, research practices. Existing methodological work on trials with adult populations should be extended to look at research with children and young people.

Further work on research synthesis methods is required to establish how a diversity perspective can best inform the conduct of reviews and secondary research.

Research governance

Research funders and commissioners, and journal editors, should require that researchers attend to the diversity of the population in the design, conduct and reporting of research studies. What this means in practice is likely to depend on the research question, the population under study, and the methods used.

5.4.2 Diversity-related goals and questions for research

To supplement the general recommendations in the preceding section, Table 5.1 presents a schematic overview of some of the diversity-related goals appropriate to each stage of the research process, and examples of the questions and conceptual issues which might arise. This overview may offer a stimulus for future reflection on how research can promote the inclusion of diverse groups.

Table 5.1: Diversity-related goals and questions for research

Research stage	Goals for research	Questions / conceptual issues
Setting research agendas and deciding on questions	<ul style="list-style-type: none"> • Involve diverse populations in identifying priority areas of research. • Use methods which are relevant and appropriate to the population under study. • (For intervention studies) Design or adapt intervention for diverse groups. 	<ul style="list-style-type: none"> • What form of involvement is appropriate (peer researchers, research assistants, action research, consultation etc.)? • Do the methods for involvement allow participants to challenge key assumptions? • Will involvement feed into the research process and help to make research more relevant and accessible to diverse groups?
Measurement and reporting of diversity	<ul style="list-style-type: none"> • Identify relevant dimensions of diversity and appropriate measures for each. 	<ul style="list-style-type: none"> • Are baseline demographics reported? • Are standardised categories or open-ended selection of dimensions more appropriate? • Are the categories and measurement tools used acceptable and relevant to participants?
Sampling	<ul style="list-style-type: none"> • Obtain a sample which reflects the diversity of the population under study. 	<ul style="list-style-type: none"> • Is statistical representativeness an appropriate goal of sampling? If so, how should it be pursued? What are the socio-demographic characteristics of the intended sample and the sample obtained? • What strategies might be employed for theoretical sampling, beyond a convenience sample? How might the inclusion of diverse perspectives lead to more rigorously supported findings? • Do the sampling methods employed lead to a sufficiently diverse sample to represent all perspectives which are relevant to the research question?

Table 5.1: Diversity-related goals and questions for research (*continued*)

Research stage	Goals for research	Questions / conceptual issues
Recruitment	<ul style="list-style-type: none"> Use recruitment methods which allow individuals from diverse groups an equal chance to participate, and do not create barriers to the participation of particular groups. 	<ul style="list-style-type: none"> What potential barriers exist to the recruitment and retention of diverse populations (language, cost, accessibility, relevance of research to minority groups, issues of trust between minority groups and researchers, etc.)? Will participants be accessed through institutional settings or gatekeepers? How might this affect recruitment rates in diverse groups? What is the process for obtaining informed consent? Is it appropriate to attempt to include hard-to-reach groups, such as children and young people not in school?
Data collection	<ul style="list-style-type: none"> Use methods of data collection which allow individuals from diverse groups to contribute to the research. 	<ul style="list-style-type: none"> Are data collection tools appropriate for the age group under study? Are data collection tools culturally appropriate for diverse populations? Is it necessary to use translators or interpreters for minority languages? Have data collection tools been validated with diverse populations? Could participants or research assistants offer useful advice on the design of data collection? Are some participants likely to contribute less than others? How can elite bias be avoided?
Data analysis	<ul style="list-style-type: none"> Analyse data so as to identify differences and commonalities between diverse groups. 	<ul style="list-style-type: none"> Are subgroup analyses and/or interaction tests appropriate tools to identify differences between groups? How can analysis seek to identify differences between diverse groups, as well as diversity within particular groups, without over-emphasising difference for its own sake? What is the rationale for undertaking particular analyses (e.g. comparing males and females, or different ethnic groups)? What is the rationale for not undertaking such analyses where they might be relevant?

6. REFERENCES

- Aggleton P, McClean C, Taylor-Laybourn A, Waller D, Warwick I, Woodhead D, Youdell D (1995) *Young Men Speaking Out*. London: Health Education Authority.
- Alderson P, Morrow V (2004) *Ethics, Social Research and Consulting with Children and Young People*. Barking: Barnardo's.
- Alderson P (2008) *Young Children's Rights: Exploring Beliefs, Principles and Practice* (2nd ed.). London: Jessica Kingsley.
- Allen I, Bourke-Dowling S, Rolfe H (1998) *Teenage Mothers: Decisions and Outcomes*. London: Policy Studies Institute.
- Allmark P (2004) Should research samples reflect the diversity of the population? *Journal of Medical Ethics* **30**: 185-189.
- Anderson AS, Hetherington M, Adamson A, Porteous L, Higgins C, Foster E, Stead M, Ha MA (2003) *The Development and Evaluation of a Novel School Based Intervention to Increase Fruit and Vegetable Intake in Children*. London: Food Standards Agency.
- Arai L (2004) *Teenage Pregnancy and Fertility in English Communities: Neighbourhood, Family and Peer Influences on Behaviour*. Unpublished doctoral thesis: University of London.
- Armstrong C, Hill M, Secker J (1998). *Listening to Children*. London: Mental Health Foundation.
- Ashcroft RE, Chadwick DW, Clark SRL, Edwards RHT, Frith L, Hutton JL (1997) Implications of socio-cultural contexts for the ethics of clinical trials. *Health Technology Assessment* **1** (9).
- Aspinall PJ (2007) Language ability: a neglected dimension in the profiling of populations and health service users *Health Education Journal* **66**: 90-106.
- Atkin K, Chattoo S (2006) Approaches to conducting qualitative research in ethnically diverse populations. In: Nazroo JY (ed.) *Health and Social Research in Multiethnic Societies*. London: Routledge, pages 95-115.
- Atkinson R, Flint J (2001). Accessing hidden and hard-to-reach populations: snowball research strategies. *Social Research Update* 33. (<http://sru.soc.surrey.ac.uk/SRU33.html>, accessed 17 September 2007).
- Badger F, Werrett J (2005) Room for improvement? Reporting response rates and recruitment in nursing research in the past decade *Journal of Advanced Nursing* **51**: 502-510.
- Baird KL (1999) The new NIH and FDA medical research policies: targeting gender, promoting justice. *Journal of Health Politics, Policy and Law* **24**: 531-565.
- Barrell M (2000) Educational experiences of adolescent mothers. *The Practising Midwife* **3**: 30-32.

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Bartlett C, Davey P, Dieppe P, Doyal L, Ebrahim S, Egger M (2003) Women, older persons and ethnic minorities: factors associated with their inclusion in randomised trials of statins 1990 to 2001. *Heart* **89**: 327-328.

Bartlett C, Doyal L, Ebrahim S, Davey P, Bachmann M, Egger M, Dieppe P (2005) The causes and effects of socio-demographic exclusions from clinical trials. *Health Technology Assessment* **9** (38).

Bendelow GA, France A, Williams S (1998) *Beliefs of Young People in Relation to Health, Risk and Lifestyles*. Coventry: Centre for Research in Health, Medicine and Society, University of Warwick.

Bhopal R, Vettini A, Hunt S, Wiebe S, Hanna L, Amos A (2004) Review of prevalence data in, and evaluation of methods for cross cultural adaptation of, UK surveys on tobacco and alcohol in ethnic minority groups. *British Medical Journal* **328**: 76.

Bickerstaff K, Shaw S (2000) *An Evaluation of the Walking Bus at Pirehill First School*. Staffordshire: The Centre for Alternative and Sustainable Transport (CAST), Staffordshire University.

Birtwistle GE, Brodie DA (1991) Children's attitudes towards activity and perceptions of physical education. *Health Education Research* **6**: 465-478.

Bloch A (2007) Methodological challenges for national and multi-sited comparative survey research. *Journal of Refugee Studies* **20**: 230-247.

Boehmer U (2002) Twenty years of public health research: inclusion of lesbian, gay, bisexual, and transgender populations. *American Journal of Public Health* **92**: 1125-1130.

Bornhöft G, Maxon-Bergemann S, Wolf U, Kienle GS, Michaelsen A, Vollman HC, Gilbertson S, Matthiessen PF (2006) Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Medical Research Methodology* 656: doi:10.1186/1471-2288-6-56.

Bostock L (2001) Pathways of disadvantage? Walking as a mode of transport among low-income mothers. *Health and Social Care in the Community* **9**: 11-18.

Boulton M, Fitzpatrick R, Swinburn C (1996) Qualitative research in health care: II. A structured review and evaluation of studies. *Journal of Evaluation in Clinical Practice* **2**: 171-179.

Boulton M, Fitzpatrick R (1997) Evaluating qualitative research. *Evidence-based Health Policy and Management* **1**: 83-85.

Bowen C (1997) School survey highlights teenage problem area. *Nursing Times* **93**: 54-55.

Britten N, Jones R, Murphy E, Stacy R (1995) Qualitative research methods in general practice and primary care. *Family Practice* **12**: 104-114.

Britton A, McKee M, Black N, McPherson K, Sanderson C, Bain C (1999) Threats to applicability of randomized trials: exclusions and selective participation. *Journal of Health Services, Research, and Policy* **4**: 112-121.

Brogan DJ, Denniston MM, Liff JM, Flagg EW, Coates RJ, Brinton LA (2001) Comparison of telephone sampling and area sampling: response rates and within-household coverage. *American Journal of Epidemiology* **153**: 1119-1127.

Brookes ST, Whitley E, Peters TJ, Mulheran PA, Egger M, Davey Smith G (2001) Subgroup analyses in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technology Assessment* **5** (33).

Brunton G, Harden A, Rees R, Kavanagh J, Oliver S, Oakley A (2003) *Children and physical activity: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Brunton G, Oliver S, Oliver K, Lorenc T (2006) *A synthesis of research addressing children's, young people's and parents' views of walking and cycling for transport*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Cahill M, Ruben T, Winn S (1996) *Children and Transport: travel patterns, attitudes and leisure activities of children in the Brighton area*. University of Brighton, Health and Social Policy Research Centre.

Campbell H, Boyd KM, Surry SAM (1997) Journals should require routine reporting of consent rates (letter). *British Medical Journal* **315**: 247.

Campbell MK, Elbourne DR, Altman DG for the CONSORT Group (2004) CONSORT statement: extension to cluster randomised trials. *British Medical Journal* **328**: 702-708.

Carr-Hill R, Passingham S, Wolf A, Kent N (1996) *Lost Opportunities: The language skills of linguistic minorities in England and Wales*. London: The Basic Skills Agency.

Chase E, Knight A, Warwick I, Aggleton P (2003) *Pregnancy and Parenthood Amongst Young People In and Leaving Care. A Draft Final Report to the Department of Health, England*. London: Thomas Coram Research Unit.

Chinn L, Elliott M, Sentinella J, Williams K (2004) *Road Safety Behaviour of Adolescent Children in Groups*. Crowthorne: Transport Research Laboratory.

Clay DL, Mordhorst MJ, Lehn L (2002) Empirically supported treatments in pediatric psychology: where is the diversity? *Journal of Pediatric Psychology* **27**: 325-337.

Coakley J, White A (1992) Making decisions: gender and sport participation among British adolescents. *Sociology of Sport Journal* **9**: 20-35.

Cocks AJ (2006) The ethical maze: finding an inclusive path towards gaining children's agreement to research participation. *Childhood* **13**: 247-266.

Coleman L, Cater S (2005) *Underage Risky Drinking*. York: Joseph Rowntree Foundation.

Corbett KK, Nettekoven L, Churchill LC, Dalton LT, Johnson CL, Dickinson L, Sorensen G, Thompson B (1995) Involving diverse community organizations in tobacco control activities. In: Thompson B, Burns DM, Lynn WR, Shopland DR (eds) *Smoking and Tobacco Control Monograph 6. Community-Based Interventions for Smokers: The COMMIT Field Experience*. Darby: Diane, pages 183-208.

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Danton K, Misselke L, Bacon R, Done J (2003) Attitudes of young people toward driving after smoking cannabis or after drinking alcohol. *Health Education Journal* **62**: 50-60.

Davis A, Jones L (1996) Environmental constraints on health: listening to children's views. *Health Education Journal* **55**: 363-374.

De Graaf R, Bijl RV, Smit F, Ravelli A, Vollebergh WAM (2000) Psychiatric and Socio-demographic Predictors of Attrition in a Longitudinal Study The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *American Journal of Epidemiology* **152**: 1039-1047.

Denscombe M (1999) Critical incidents and invulnerability to risk: young people's experience of serious health-related incidents and their willingness to take health risks. *Health, Risk and Society* **1**: 195-207.

Denzin NK, Lincoln YS (eds) (2005) *The Sage Handbook of Qualitative Research* (3rd ed.). Thousand Oaks: Sage.

Department for Communities and Local Government (2007) *A Framework for Fairness: Proposals for a Single Equality Bill for Great Britain - A Consultation Paper*. Wetherby: Communities and Local Government Publications.
(<http://www.communities.gov.uk/documents/corporate/pdf/325332>, accessed 6 September 2007).

Department of Health (2003) *Addressing the Race Relations Amendment Act: Guidelines for Commissioners in the Department of Health Policy Research Programme. Fifth draft: June 2003*. Unpublished report.

Department of Health (2005) *Research Governance Framework for Health and Social Care*. London: Department of Health.

Department of Health (2007) *Single Equality Scheme 2007-2010*. London: Department of Health.

De Vaus D (2002) *Surveys in Social Research*. London: Routledge.

Dieppe P, Bartlett C, Davey P, Doyal L, Ebrahim S (2004) Balancing benefits and harms: the example of non-steroidal anti-inflammatory drugs. *British Medical Journal* **329**: 31-34.

Dixon-Woods M, Cavers D, Agarwal S, Annandale E, Arthur A, Harvey J, Hsu R, Katbamna S, Olsen R, Smith L, Riley R, Sutton AJ (2006) Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology* **6**: 35.

Dixon-Woods M, Shaw RL, Agarwal S, Smith JA (2004) The problem of appraising qualitative research. *Quality and Safety in Health Care* **13**: 223-225.

Edwards SJL, Lilford RJ, Braunholtz DA, Jackson JC, Hewison J, Thornton J (1998) Ethical issues in the design and conduct of randomised controlled trials. *Health Technology Assessment* **2** (15).

Elliot E, Watson A, Harries U (2002) Harnessing expertise: involving peer interviewers in qualitative research with hard-to-reach populations. *Health Expectations* **5**: 172-178.

- Emmel N, Hughes K, Greenhalgh J, Sales A (2007) Accessing socially excluded people: trust and the gatekeeper in the researcher-participant relationship. *Sociological Research Online* **12** (2).
- Engineer R, Phillips A, Thompson J, Nicholls J (2003) *Drunk and Disorderly: A Qualitative Study of Binge Drinking Among 18-to-24-year-olds*. London: Home Office.
- Epstein LH, Gordy CC, Raynor HA, Beddome M, Kilanowski CK, Paluch R (2001) Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity. *Obesity Research* **9**: 171-178.
- Estabrooks P, Dzewaltowski DA, Glasgow RE, Klesges LM (2003) Reporting of validity from school health promotion studies published in 12 leading journals, 1996-2000. *Journal of School Health* **73**: 21-28.
- Evans T, Brown H (2003) Road traffic crashes: operationalizing equity in the context of health sector reform. *Injury Control and Safety Promotion* **10**: 11-12.
- Faugier J, Sargeant M (1997) Sampling hard to reach populations. *Journal of Advanced Nursing* **26**: 790-797.
- Flaskerud JH, Nyamathi AM (2000) Attaining gender and ethnic diversity in health intervention research: cultural responsiveness versus resource provision. *Advances in Nursing Science* **22**: 1-15.
- Flick U (2006) *An Introduction to Qualitative Research* (3rd ed.). London: Sage.
- Fowler FJ (2002) *Survey Research Methods* (3rd ed.). Thousand Oaks: Sage.
- Fraser N, Lewis V, Ding S, Kellett M, Robinson C (eds) (2004) *Doing Research with Children and Young People*. London: Sage.
- Fredman S, Spencer S (2006) *Delivering Equality: Towards an Outcome-Focused Positive Duty*. Submission to the Cabinet Office Equality Review and to the Discrimination Law Review, June 2006. (<http://www.edf.org.uk/news/Delivering%20equality%20submission%20030606-final.pdf>, accessed 6 September 2007).
- Gallagher M, Millar R, Hargie O, Ellis R (1992) The personal and social worries of adolescents in Northern Ireland: results of a survey. *British Journal of Guidance and Counselling* **20**: 274-290.
- Gibson EL, Wardle J, Watts CJ (1998) Fruit and vegetable consumption, nutritional knowledge and beliefs in mothers and children. *Appetite* **31**: 205-228.
- Gilson SF, DePoy E (2004) Disability, identity and cultural diversity. *Review of Disability Studies* **1**: 16-23.
- Glaser BG, Strauss AL (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Glasgow RE, Bull SS, Gillette C, Klesges LM, Dzewaltowski DA (2002) Behavior change intervention research in healthcare settings: a review of recent reports with emphasis on external validity. *American Journal of Preventive Medicine* **24**: 62-69.

Gortmaker SL, Cheung LW, Peterson KE, Chomitz G, Cradle JH, Dart H, Fox MK, Bullock RB, Sobol AM, Colditz G, Field AE, Laird N (1999) Impact of a school-based interdisciplinary intervention on diet and physical activity among urban primary school children: eat well and keep moving. *Archives of Pediatrics and Adolescent Medicine* **153**: 975-983.

Graham H (2006) Social determinants and public health policy in the UK. In Killoran A, Swann, C, Kelly M (eds) *An evidence-based approach to public health and tackling health inequalities: opportunities and challenges*. Oxford: Oxford University Press.

Greenhalgh T, Peacock R (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *British Medical Journal* **331**: 1064-1065.

Greenwood DJ, Levin M (1998) *Introduction to Action Research: Social Research for Social Change*. Thousand Oaks: Sage.

Grewal I, Ritchie J (2006) Ethnic and language matching of the researcher and the research group during design, fieldwork and analysis. In: Nazroo JY (ed) *Health and Social Research in Multiethnic Societies*. London: Routledge, pages 65-81.

Gross CP, Mallory R, Helat A, Krumholz HM (2002) Reporting the recruitment process in clinical trials; who are these patients and how did they get there? *Annals of Internal Medicine* **137**: E-10-E-16.

Gunaratnam Y (2003) *Researching Race and Ethnicity: Methods, Knowledge and Power*. London: Sage.

Hahn A, Leavitt T, Aaron P (1994) *Evaluation of the Quantum Opportunities Program (QOP): Did the Program Work? A Report of the Post-Secondary Outcomes and Cost-Effectiveness of the QOP (1989-1993)*. Waltham, Massachusetts, USA: Brandeis University, Heller Graduate School.

Hall D, Rae C, Jarvis M (2003) *Young Parent Peer Research Project*. Leicester: Connexions Leicestershire.

Hall WD (1999) Representation of blacks, women, and the very elderly (aged ≥ 80) in 28 major randomized clinical trials. *Ethnicity and Disease* **9**: 333-340.

Harden A, Rees R, Shepherd J, Brunton G, Oliver S, Oakley A (2001) *Young People and Mental Health: a systematic review of research on barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Harden A, Brunton G, Fletcher A, Oakley A, Burchett H, Backhans M (2006) *Young People, Pregnancy and Social Exclusion: a systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Harris J (1994) Young people's perceptions of health, fitness and exercise: implications for the teaching of health related exercise. *Physical Education Review* **17**: 143-151.

- Hart KH, Bishop JA, Truby H (2002) An investigation into school children's knowledge and awareness of food and nutrition. *Journal of Human Nutrition and Dietetics* **15**: 129-140.
- Hatch JA (ed) (1995) *Qualitative Research in Early Childhood Settings*. London: Praeger.
- Hille ETM, den Ouden AL, Stuijbergen MC, Verrips GHW, Vogels AGC, Brand R, Bennebroek Gravenhorst J, Verloove-Vanhorick SP (2005) Is attrition bias a problem in neonatal follow-up? *Early Human Development* **81**: 901-908.
- Honess T, Seymour L, Webster R (2000) *The Social Contexts of Underage Drinking*. London: Home Office.
- Hughes H, Cragg A, Taylor C (1999) *Reducing the Rate of Teenage Conceptions. Young People's Experiences of Relationships, Sex and Early Parenthood: Qualitative Research*. London: Health Education Authority.
- Hussain-Gambles M, Leese B, Atkin K, Brown J, Mason S, Tovey P (2004) Involving South Asian patients in clinical trials. *Health Technology Assessment* **8** (42).
- Johnson MRD (2006) Engaging communities and users: health and social care research with ethnic minority communities. In: Nazroo JY (ed.) *Health and Social Research in Multiethnic Societies*. London: Routledge, pages 48-64.
- Jones A (2004) Involving children and young people as researchers. In: Fraser N, Lewis V, Ding S, Kellett M, Robinson C (eds) *Doing Research with Children and Young People*. London: Sage, pages 113-130.
- Jones L (1998) Inequality in access to local environments: the experiences of Asian and non-Asian girls. *Health Education Journal* **57**: 313-328.
- Jones L, Allebone B (1999) Researching 'hard-to-reach' groups: the crucial role of the research associate. *International Journal of Inclusive Education* **3**: 353-362.
- Kahn JP, Mastroianni AC, Sugarman J (1998) *Beyond Consent: Seeking Justice in Research*. New York: Oxford University Press.
- Kalsbeek WD (2003) Sampling minority groups in health surveys. *Statistics in Medicine* **22**: 1527-1549.
- Karlsen S, Nazroo JY (2002) Agency and structure: the impact of ethnic identity and racism on the health of ethnic minority people. *Sociology of Health and Illness* **24**: 1-20.
- Kemmis S, McTaggart R (2005) Participatory action research: communicative action and the public sphere. In: Denzin NK, Lincoln YS (eds) *The Sage Handbook of Qualitative Research* (3rd ed). Thousand Oaks: Sage, pages 559-604.
- Klepp KI (1995) Nonresponse bias due to consent procedures in school-based, health-related research. *Scandinavian Journal of Social Medicine* **23**: 53-59.
- Knight A, Clark A, Petrie P, Statham J (2006) *The Views of Children and Young People with Learning Disabilities about the Support they Receive from Social Services: A Review of*

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Consultations and Methods. London: Thomas Coram Research Unit, Institute of Education, University of London.

Lee A (1993). Peer pressure in cycle helmet use. *Paediatric Nursing* **5**: 12-15.

Levine RJ (1991) Informed consent: some challenges to the universal validity of the Western model. *Law, Medicine and Health Care* **19**: 207-213.

Lincoln YS, Guba EG (1985) *Naturalistic Inquiry*. London: Sage.

Lippman A (2006) *The Inclusion of Women in Clinical Trials: Are We Asking the Right Questions?* Toronto: Women and Health Protection.

Liquori T, Koch PD, Contento IR, Castle J (1998) The Cookshop Program: outcome evaluation of a nutrition education program linking lunchroom food experiences with classroom cooking experiences. *Journal of Nutrition Education* **30**: 302-313.

Long AF, Godfrey M (2004) An evaluation tool to assess the quality of qualitative research studies. *International Journal of Social Research Methodology* **7**: 181-196.

Luepker RV, Perry CL, McKinlay SM, Nader PR, Parcel GS, Stone EJ, Webber LS, Elder JP, Feldman HA, Johnson CC, Kelder SH, Wu M (1996) Outcomes of a field trial to improve children's dietary patterns and physical activity. The Child and Adolescent Trial for Cardiovascular Health (CATCH). *Journal of the American Medical Association* **275**: 768-776.

MacNaughton G, Rolfe SA, Siraj-Blatchford I (eds) (2001) *Doing Early Childhood Research: International Perspectives on Theory and Practice*. Buckingham: Open University Press.

Magnani R, Sabin K, Saidel T, Heckathorn D (2005) Review of sampling hard-to-reach and hidden populations for HIV surveillance. *AIDS* **19** (suppl 2): S67-S72.

Mason S, Hussain-Gambles M, Leese B, Atkin K, Brown J (2003) Representation of South Asian people in randomised clinical trials: analysis of trials' data. *British Medical Journal* **326**: 1244-1245.

Mauthner M, Mayall B, Turner S (1993) *Children and Food at Primary School*. London: Social Science Research Unit, Institute of Education, University of London.

McCrystal P, Percy A, Higgins K (2007) Exclusion and marginalisation in adolescence: the experience of school exclusion on drug use and antisocial behaviour. *Journal of Youth Studies* **10**: 35-54.

Miller K, McKeever P, Coyte P (2003) Recruitment issues in healthcare research: the situation in home care. *Health and Social Care in the Community* **11**: 111-123.

Mitchell K (1997) Encouraging young women to exercise: can teenage magazines play a role? *Health Education Journal* **56**: 264-273.

Moher D, Schulz KM, Altman DG (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* **357**: 1191-1194.

- Moore L (2001) *Are Fruit Tuck Shops in Primary Schools Effective in Increasing Pupils' Fruit Consumption? A Randomised Controlled Trial*. Cardiff: Cardiff School of Social Sciences.
- Morrow V, Richards M (1996) The ethics of social research with children: an overview. *Children and Society* **10**: 90-105.
- Mulvihill C, Rivers K, Aggleton P (2000) *Physical Activity 'At Our Time'*. London: Health Education Authority.
- National Co-ordinating Centre for Research Methodology (2001) *Addressing Ethnic Diversity in Health and Social Care Research*. Birmingham: University of Birmingham.
- National Institutes of Health (2001) *NIH Policy and Guidelines on the Inclusion of Women and Minorities as Subjects in Clinical Research – amended October 2001*. http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm (accessed 6 September 2007).
- National Statistics (2007a) *Pupil Absence in Secondary Schools in England, 2005/06 (SFR 11/2007)*. London: Department for Education and Skills. (<http://www.dfes.gov.uk/rsgateway/DB/SFR/s000718/SFR11-2007.pdf>, accessed 6 September 2007).
- National Statistics (2007b) *Permanent and Fixed Period Exclusions from Schools and Exclusion Appeals in England (SFR 21/2007)*. London: Department for Education and Skills. (<http://www.dfes.gov.uk/rsgateway/DB/SFR/s000733/SFR21-2007.pdf>, accessed 6 September 2007).
- Nazroo JY (ed.) (2006) *Health and Social Research in Multiethnic Societies*. London: Routledge.
- Neale RJ, Otte S, Tilston CH (1998) Fruit: comparisons of attitudes, knowledge and preferences of primary school children in England and Germany. *Zeitschrift für Ernährungswissenschaft* **37** (suppl. 1): 123-130.
- Ness RB, Nelson DB, Kumanyika SK, Grisso JA (1997) Evaluating minority recruitment into clinical studies: how good are the data? *Annals of Epidemiology* **7**: 472-478.
- Newman M, Elbourne D (2004) Improving the usability of educational research: guidelines for REPORTing of Primary Empirical Research Studies in Education (The REPOSE Guidelines). *Evaluation and Research in Education* **18**: 201-212.
- Nickens HW (1990) Health promotion and disease prevention among minorities. *Health Affairs* **9**: 133-143.
- Nicklas TA, Johnson CC, Myers L, Farris RP, Cunningham A (1998) Outcomes of a high school programme to increase fruit and vegetable consumption: Gimme 5 – a fresh nutrition concept for students. *Journal of School Health* **68**: 248-253.
- Oakley A (2006) Ethnicity and research evaluating health interventions: issues of science and ethics. In: Nazroo JY (ed.) *Health and Social Research in Multiethnic Societies*. London: Routledge, pages 142-164.

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Oakley A, Peersman G, Oliver S (1998) Social characteristics of participants in health promotion effectiveness research: trial and error? *Education for Health* **11**: 305-317.

Oakley A, Wiggins M, Turner H, Rajan L, Barker M (2003) Including culturally diverse samples in health research: a case study of an urban trial of social support. *Ethnicity and Health* **8**: 29-39.

Oliver S, Kavanagh J, Caird J, Lorenc T, Oliver K, Harden A, Thomas J, Greaves A, Oakley A (2008) *Health promotion, inequalities and young people's health: a systematic review of research*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Onwuegbuzie AJ, Leech NL (2007) A call for qualitative power analyses. *Quality and Quantity* **41**: 105-121.

Osborn DPJ (1999) Research and ethics: leaving exclusion behind. *Current Opinion in Psychiatry* **12**: 601-604.

Parkes J, Kerr C, McDowell B, Cosgrove A (2006) Recruitment bias in a population-based study of children with cerebral palsy. *Pediatrics* **118**: 1616-1622.

Peersman G (2001) Promotion health: principles of practice and evaluation. In: Oliver S, Peersman G (eds) *Using Research for Effective Health Promotion*. Buckingham: Open University.

Perry CL, Klepp KI, Dudovitz B, Golden D, Griffin, Smyth M (1987) Promoting healthy eating and physical activity patterns among adolescents: a pilot study of 'Slice of Life'. *Health Education Quarterly* **2**: 93-103.

Perry CL, Lytle LA, Feldman H, Nicklas T, Stone E, Zive M, Garceau A, Kelder SH (1998) Effects of the child and adolescent trial for cardiovascular health (CATCH) of fruit and vegetable intake. *Journal of Nutrition Education* **30**: 354-360.

Phoenix A (1994) Practising feminist research: the intersection of gender and 'race' in the research process. In: Maynard M and Purvis J (eds) *Researching Women's Lives from a Feminist Perspective*. London: Taylor and Francis, pages 49-71.

Prescott RJ, Counsell CE, Gillespie WJ, Grant AM, Russell IT, Kiauka S, Colthart IR, Ross S, Shepherd SM, Russell D (1999) Factors that limit the quality, number and progress of randomised controlled trials. *Health Technology Assessment* **3** (20).

Rees R, Harden A, Shepherd J, Brunton G, Oliver S, Oakley A (2001) *Young People and Physical Activity: A Systematic Review of Research on Barriers and Facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Rees RW, Garcia J, Oakley A (2007) Consent in school-based research involving children and young people: a survey of research from systematic reviews. *Research Ethics Review* **3**: 35-39.

Rogers A, Adamson JE, McCarthy M (1997) Variations in health behaviours among inner city 12-year-olds from four ethnic groups. *Ethnicity and Health* **2**: 302-316.

- Rothwell PM (2005a) External validity of randomised controlled trials: 'To whom do the results of this trial apply?'. *Lancet* **365**: 82-93.
- Rothwell PM (2005b) Subgroup analysis in randomised controlled trials: importance, indications and interpretation. *Lancet* **365**: 176-186.
- Sahota P, Rudolf MCJ, Dixey R, Hill AJ, Barth JH, Cade J (2001) Randomised controlled trial of primary school based intervention to reduce risk factors for obesity. *British Medical Journal* **323**: 1029-1032.
- Sawyer L, Regev H, Proctor S, Nelson M, Messias D, Barnes D, Meleis AI (1995) Matching versus cultural competence in research: methodological considerations. *Research in Nursing and Health* **18**: 557-567.
- Schünemann HJ, Fretheim A, Oxman AD (2006) Improving the use of research evidence in guideline development: 10. Integrating values and consumer involvement. *Health Research Policy and Systems* **4**: 22.
- Scott Porter Research and Marketing (2000) *Young People and Mental Well Being*. Edinburgh: Health Education for Scotland.
- Scottish Executive (2006) *Children Educated Outwith School and Pupil Projections, 2006*. Edinburgh: Scottish Executive.
<https://www.scotland.gov.uk/Publications/2006/11/10144033/12> (accessed 6 September 2007).
- Sheikh A, Netuveli G, Kai J, Panesar SS (2004) Comparison of reporting of ethnicity in US and European randomised controlled trials. *British Medical Journal* **329**: 87-88.
- Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, Oakley A (2001) *Young people and healthy eating: a systematic review of research on barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Siddiqui O, Flay BR, Hu FB (1996) Factors affecting attrition in a longitudinal smoking prevention study. *Preventive Medicine* **25**: 554-560.
- Siegel JM, Yancey AK, Aneshensel CS, Schuler R (1999) Body image, perceived pubertal timing, and adolescent mental health. *Journal of Adolescent Health* **25**: 155-165.
- Silagy C, Jewell D (1994) Review of 39 years of randomized controlled trials in the British Journal of General Practice. *British Journal of General Practice* **44**: 359-363.
- Sin CH (2004) Sampling minority ethnic older people in Britain. *Ageing and Society* **24**: 257-277.
- Speak S (1995) *Young Single Mothers: Barriers to Independent Living*. Oxford: Family Policy Studies Centre.
- Spencer L, Ritchie J, Lewis J, Dillon L (2003) *Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence*. London: Cabinet Office.

Including diverse groups of children and young people in health promotion and public health research: a review of methodology and practice

Swanson GM, Ward AJ (1995) Recruiting minorities into clinical trials: toward a participant-friendly system. *Journal of the National Cancer Institute* **87**: 1747-1759.

Taylor SB, Halliday ME (1996) *Cycle Helmet Wearing in Great Britain*. Crowthorne: Transport Research Laboratory.

Temple B (2002) Crossed wires: Interpreters, translators and bilingual workers in cross language research. *Qualitative Health Research* **12**: 844-854.

Temple B (2005) Nice and tidy: translation and representation. *Sociological Research Online* **10** (2).

Testa AC, Coleman LM (2006) Accessing research participants in schools: a case study of a UK adolescent sexual health survey. *Health Education Research* **21**: 518-526.

Thomas J, Sutcliffe K, Harden A, Oakley A, Oliver S, Rees R, Brunton G, Kavanagh J (2003) *Children and healthy eating: a systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Thomas J, Kavanagh J, Tucker H, Burchett H, Tripney J, Oakley A (2007) *Accidental injury, risk-taking behaviour and the social circumstances in which young people live: a systematic review*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Tong A, Sainsbury P, Craig J (2007) Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care* doi:10.1093/intqhc/mzm042.

Tyrrell H (2000) *Going My Way: What Children and Young People Say about Transport*. Edinburgh: Save the Children Scotland.

Unger JB, Gallaher P, Palmer PH, Baezconde-Garbanati L, Trinidad DL, Cen S, Anderson Johnson C (2004) No news is bad news: characteristics of adolescents who provide neither parental consent nor refusal for participation in school-based survey research. *Evaluation Review* **28**: 52-63.

Van Spall HGC, Toren A, Kiss A, Fowler RA (2007) Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *Journal of the American Medical Association* **297**: 1233-1240.

Waldenström U, Turnbull D (1998) A systematic review comparing continuity of midwifery care with standard maternity services. *British Journal of Obstetrics and Gynaecology* **105**:1160-1170.

Wolf HK, Kuulasmaa K, Tolonen H, Sans S, Molarius A, Eastwood BJ (2005) Effect of sampling frames on response rates in the WHO MONICA risk factor surveys. *European Journal of Epidemiology* **20**: 293-299.

Yancey AK, Ortega AN, Kumanyika SK (2006) Effective recruitment and retention of minority research participants. *Annual Review of Public Health* **27**: 1-28.

First produced in 2008 by:

Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre)
Social Science Research Unit
Institute of Education, University of London
18 Woburn Square
London WC1H 0NR

Tel: +44 (0)20 7612 6367

<http://eppi.ioe.ac.uk/>

<http://www.ioe.ac.uk/ssru/>

ISBN: 978-0-9559087-4-3

The **Evidence for Policy and Practice Information and Co-ordinating Centre** (EPPI-Centre) is part of the Social Science Research Unit (SSRU), Institute of Education, University of London.

The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the Institute of Education, University of London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

The views expressed in this work are those of the authors and do not necessarily reflect the views of the EPPI-Centre or the funder. All errors and omissions remain those of the authors.

This document is available in a range of accessible formats including large print. Please contact the Institute of Education for assistance:
telephone: +44 (0)20 7947 9556 email: info@ioe.ac.uk